



INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

I. A. R. I. C.

MGIPC—84—10 AR—21-6 49—1,000.

PROCEEDINGS
OF THE
ROYAL SOCIETY OF LONDON

SERIES A. MATHEMATICAL AND PHYSICAL SCIENCES

VOL CLXVIII

LONDON

Printed and published for the Royal Society
By the Cambridge University Press
Bentley House, N.W. 1

25 November 1938

PRINTED IN GREAT BRITAIN BY
WALTER LEWIS, M.A.
AT THE CAMBRIDGE UNIVERSITY PRESS

CONTENTS

SERIES A VOL CLXVIII

No. A 932. 10 October 1938

	PAGE
The decomposition and ignition of peroxides. I. Diethylperoxide. By E. J. Harris and A. C. Egerton, F.R.S.	1
Formation of negative ions at metal surfaces. By R. A. Smith . . .	19
Thermal effects on bodies in an air stream. By W. F. Hilton . . .	43
Statistical theory of superlattices with long-range interaction. I. General theory. By J. S. Wang	56
Statistical theory of superlattices with long-range interaction. II. The simple cubic lattice and the body-centred cubic lattice. By J. S. Wang	68
The nuclear spin of iodine. III. Further measurements upon the fine structures in the first spark spectrum. By S. Tolansky and G. O. Forester. (Plate 1)	78
A new process of negative-ion formation. IV. By F. L. Arnot and Clark Beckett	103
New systems of normal co-ordinates for relativistic optics. By G. Temple	122
The separation of isotopes for the investigation of nuclear transmutations. By E. Leighton Yates	148

No. A 933. 25 October 1938

The scattering of fast β -particles. By A. Barber and F. C. Champion. (Plate 2)	159
Photographic sensitivity and the reciprocity law at low temperatures. By W. F. Berg and K. Mendelssohn	168
Experiments on the transmutation of fluorine by protons and deuterons. By W. E. Burcham and C. L. Smith	176
The thermal properties and heats of adsorption of films on vitreous silica. By W. G. Palmer	190
The rates of transformation in ethyl alcohol of ammonium and ethyl-ammonium cyanates to the corresponding azides. By C. C. Miller and J. R. Nicholson	206

	PAGE
The use of a vertical pipe as an overflow for a large tank. By A. M. Binnie. (Plates 3-6)	219
Specific heat-temperature curves of some age-hardening alloys. By N. Swindells and C. Sykes	237
The classical equation of state of gaseous helium, neon and argon. By R. A. Buckingham	264
The formation of negative ions by positive-ion impact on surfaces. By R. H. Sloane and R. Press	284

No. A 934. 7 November 1938

A discussion on plastic flow in metals. By W. L. Bragg, F.R.S.	302
Relaxation methods applied to engineering problems. III. Problems involving two independent variables. By D. G. Christopherson and R. V. Southwell, F.R.S.	317
Oscillatory motion of a fluid along a circular tube. By D. G. Christopherson, A. Gemant, A. H. A. Hogg and R. V. Southwell, F.R.S.	351
The low-temperature properties of gaseous helium. By H. S. W. Massey and R. A. Buckingham	378
The electromagnetic energy of a point charge. By M. H. L. Pryce	389
The porous diaphragm method of measuring diffusion velocity, and the velocity of diffusion of potassium chloride in water. By G. S. Hartley and D. F. Runnicles	401
The determination of the size of paraffin-chain salt micelles from diffusion measurements. By G. S. Hartley and D. F. Runnicles	420

No. A 935. 25 November 1938

The thermal decomposition of nitrous oxide. By R. M. Lewis and C. N. Hinshelwood, F.R.S.	441
Progressive lightning. VI. By B. F. J. Schonland, F.R.S., D. J. Malan and H. Collens. (Plates 7-12)	455
A new Coriolis perturbation in the methane spectrum. I. Vibrational-rotational Hamiltonian and wave functions. By H. A. Jahn	469

The decomposition and ignition of peroxides.

I. Diethylperoxide

BY E. J. HARRIS AND A. C. EGERTON, F.R.S.

(Received 24 June 1938)

INTRODUCTION

Organic peroxides are formed during the combustion of hydrocarbons (Callendar 1927; Egerton 1927; Mondain Monval and Quanquin 1929), and they have been shown to be very powerful "proknock" substances (Egerton, Smith and Ubbelohde 1935). The work of Townend (1933-8) has shown that the lower ignition region of the hydrocarbons is that which characterizes their "knock" behaviour. In this region there appear to be substances formed which are only stable within a certain range of temperature and it is those substances which seem to determine the ignition and knock behaviour (Egerton 1928; Neumann and Aivazov 1935).

It is quite likely that the substances are peroxides, but as it is not easy to decide from estimation of the products of combustion what peroxides are formed (Pease 1934, 1935; Harris and Egerton 1937) and which are important in the processes which eventually lead to ignition or give rise to knock, it is necessary to obtain information from several directions in order to find out what exactly is happening. Experiments are therefore being undertaken to find out the behaviour of various organic peroxides when heated and when oxidized, and the present communication deals with diethylperoxide. It is hardly likely that this plays a prominent part in the combustion of a hydrocarbon and it would be easy to detect if it was formed in any quantity, but being a strong proknock, and being stable and fairly easily prepared in a pure state, it forms a convenient peroxide to study first.

Some information already exists regarding the products into which peroxides of various kinds decompose (Rieche 1931), but little is known as to the kinetics of their decomposition, though Medvedeff and Podjapolskaya (1935) have studied the decomposition of methylhydrogen-peroxide by a flow method and their interpretation of the results of their experiments indicates that a unimolecular decomposition to formaldehyde and water occurs alongside a bimolecular decomposition to alcohol and oxygen.

Diethylperoxide only decomposes very slowly below 120° C., but in the region 130–190° C the rate of decomposition is measurable.

In what follows the decomposition is shown to be mainly homogeneous and unimolecular, but like azomethane, above a certain critical pressure the decomposition becomes explosive and luminous. Neumann and Toutakin (1937) have already drawn attention to these limiting explosion pressures of diethylperoxide and have stated that butane oxygen mixtures containing diethylperoxide ignite and give rise to cool flames at the same limiting partial pressure of peroxide as would lead to explosion of the latter alone, but we do not find this behaviour in the case of mixtures with propane.

EXPERIMENTAL

(a) *Preparation*

Diethylperoxide was prepared from diethylsulphate and hydrogen peroxide (Baeyer and Villiger 1900). 120 c.c. 15 % H_2O_2 solution and 100 c.c. of solution of caustic potash (containing 50 g. KOH) are dropped separately into 200 g. diethylsulphate over a period of 2 hr. while cooled and stirred. The mixture is stirred and cooled overnight. By adopting this procedure hydrolysis of ethylhydrogen peroxide, which is formed intermediately, is diminished. The liquid which is separated contains ether, diethylperoxide and excess of ethylsulphate. After several successive distillations under reduced pressure nearly pure diethylperoxide is obtained. Boiling point, 58–59° C (at 62.5 cm. pressure); refractivity (N_{D}), 1.3758 at 9.5° C and 1.3715 at 16.5° C; analysis, C 52.8 %, H 11.5 % (theory, 53.3 and 11.1).

(b) *Thermal decomposition*

The thermal decomposition was investigated in two vessels, one a 22.4 c.c. Pyrex tube measuring 1.4 cm. diameter \times 15 cm. long, the other a 100 c.c. cylindrical quartz vessel with plane ends (4.5 cm. diameter \times 6.5 cm. long). Fig. 1 shows diagrammatically the experimental arrangement with the wider tube. The vessels were surrounded with a double-jacketed welded iron vessel containing vapours of boiling liquids such as xylene (140° C), bromobenzene (156° C), cyclohexanol acetate (175° C), decalin (192° C), mercury (357° C). The pressures were adjusted when intermediate temperatures were required, using a constant-pressure device. The temperature in the case of the large vessel was measured by insertion of a calibrated Anschütz thermometer inside the reaction vessel as shown, as well as in the vapour

in the jacket. The bulb *A* is made of such a volume that the initial pressure required in the vessel *B* is obtained conveniently; the measurement of the initial pressure is made in the vessel *B* by means of the gauge *D*; this is possible when more than a second's lag occurs before decomposition starts; if the lag is shorter the initial pressure has to be calculated from the measurements on the gauge *C*, knowing the volume of the bulb *A*.

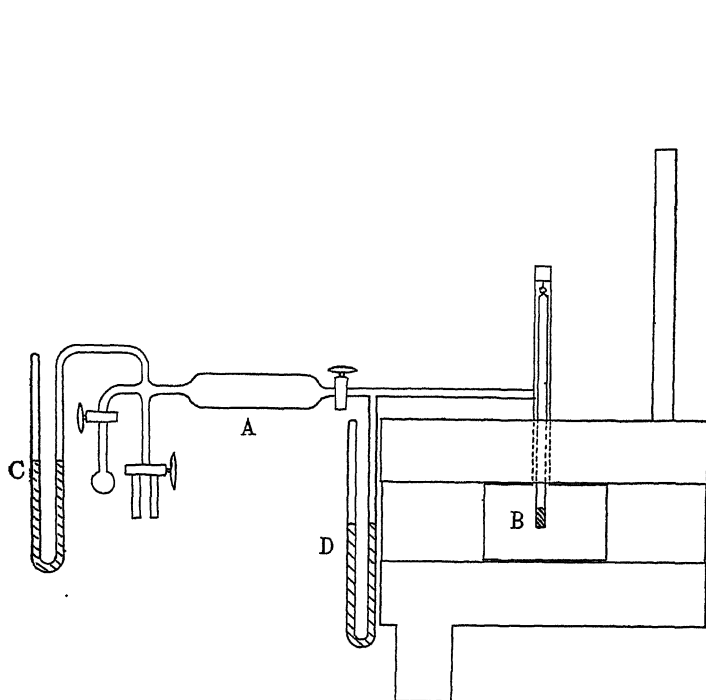


FIG. 1

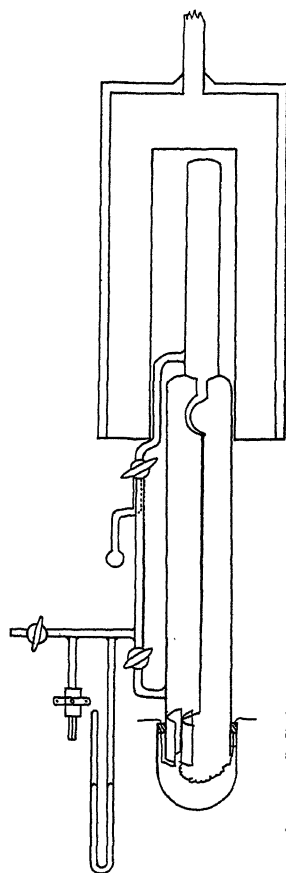


FIG. 2

In some of the later experiments on the slow decomposition, the pressure has been measured by means of a Pyrex spoon gauge (see fig. 2), the pointer of which carries a zinc plate in proximity to a fixed plate, the change of capacity due to movement of the former being measured, through a valve circuit, direct on the scale of a milliammeter. (Details of this gauge will be published elsewhere.)

EXPERIMENTAL RESULTS

The unimolecular decomposition

It is unnecessary to record all the experiments in detail. The straight line (fig. 3) drawn through the points representing thirteen different runs at 156.5°C indicates the nature of the agreement: $\log_{10} a/(a-x)$ is plotted against time, a being the initial pressure. Similar graphs are given in fig. 4

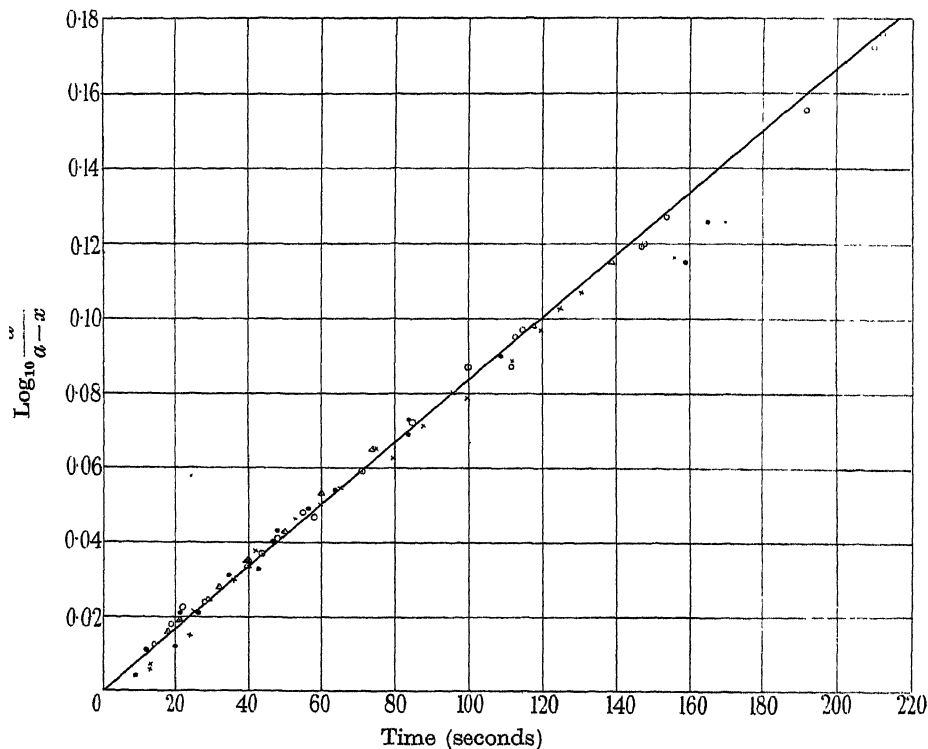


FIG. 3. Slow decomposition of diethyl peroxide at 156.5°C (13 experiments).
 \times SiO_2 ; \odot apparatus 1; \bullet apparatus 3; \triangle apparatus 4.

for other temperatures, but only selected points are shown, the number of experiments carried out at the various temperatures is given in brackets; the points plotted are generally the observations for the highest and lowest initial pressures. The initial pressures varied from about 30 to 2 cm. at various temperatures. From twenty separate experiments at temperatures from 143 to 182°C taken to the conclusion of pressure change, the final pressure is found to be $(2.17 \pm 0.03) a$ and x calculated accordingly from the observed pressure change. The graphs show that the reaction behaves unimolecularly.

In Table I are set out the initial pressures, and the apparatus used in the several experiments. Apparatus 1 refers to the small Pyrex tube, apparatus 3 to the quartz vessel, and 4 to the quartz tube and spoon manometer (4* indicates a specially sensitive spoon manometer). The amount of diluent is given and also whether the surface was increased by the addition of quartz (SiO_2) or altered by washing with salt (NaCl).

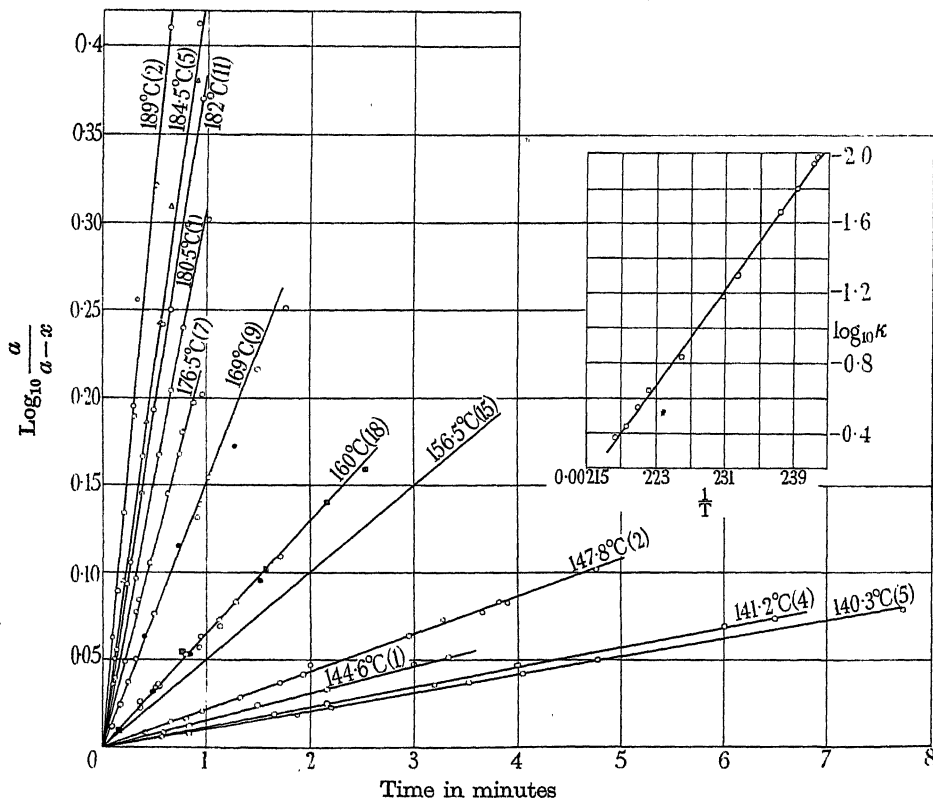


FIG. 4. Slow decomposition of diethyl peroxide. ● Helium; △ hydrogen; ▢ carbon dioxide.

Plotting $\log_{10} K$ against $1/T$, a straight line is obtained (fig. 4); by the method of least squares, the line of closest fit to the observed values is found to be $\log_{10} K = -\frac{6890.4}{T} + 14.708$; the apparent activation energy, 31,500 cal. (see Table II).

A small but erratic induction period was noticed before commencement of reaction amounting to 7 sec. at 160°C , to not more than 2 sec. at 180°C . This seemed to be related to the purity of the peroxide, for with the earlier, less pure, samples the induction was more noticeable (about 30 sec. at

TABLE I

Temp. ° C	Initial pressure cm.	Pressure of diluent	Apparatus	Temp. ° C	Initial pressure cm.	Pressure of diluent	Apparatus
140.2	32.55	—	1	160.0	3.27	+ 1.51 Air	3
140.2	24.7	—	1	160.0	3.28	—	3
140.3	15.15	—	1	160.0	2.7	—	3
140.3	8.95	—	1	160.0	2.67	+ 4.2 Air	3
140.3	5.30	—	1	160.5	6.5	—	3
141.2	6.91	—	4	160.5	4.45	—	3
141.2	6.89	—	4	169.0	6.35	—	3
141.2	6.81	—	4	169.0	3.9	—	3
141.2	4.14	—	4	169.0	3.43	—	3
144.4	9.4	—	3	169.0	1.8	+ 2.1 He	3
147.1	9.45	—	3	169.0	2.29	+ 0.4 NO	3
147.1	4.95	—	3	169.0	1.75	+ 0.35 NO	3
155	2.67	—	4 (NaCl)	169.0	1.8	+ 0.80 NO	3
156	3.0	—	4 (NaCl)	169.4	3.85	—	3
156	2.94	—	4 (NaCl)	169.4	3.85	—	3
156	2.80	—	4 (NaCl)	169.4	2.7	—	3
156.2	29.55	—	1	169.5	2.5	—	3
156.2	26.6	—	1	169.6	3.65	—	3
156.2	18.95	—	1	176.4	3.15	—	3
156.2	9.45	—	1	176.4	2.95	—	3
156.2	4.25	—	1	176.4	2.0	—	3
156.5	10.2	—	3 (SiO ₂)	176.8	3.2	—	3
156.5	9.6	+ 9.6 Air	3	176.8	3.12	—	3
156.5	8.6	—	3	176.8	2.8	—	3
156.5	7.3	—	3	176.8	2.7	—	3
156.5	5.5	—	3 (SiO ₂)	180.5	2.4	—	4 (NaCl)
156.5	5.17	+ 1.2 NO	3	181.4	2.25	—	3
156.5	2.2	—	3	181.5	0.53	—	4*
156.5	2.07	+ 2.8 Air	3	182.0	5.75	—	4
157.0	6.2	—	3 (SiO ₂)	182.0	3.7	—	4
157.0	2.3	—	3	182.0	3.55	—	4
160.0	8.9	—	3	182.0	3.04	—	4
160.0	8.9	—	3	182.0	2.92	+ 5.4 He	3
160.0	7.15	—	3	182.0	2.25	—	4
160.0	7.03	—	3	182.0	2.08	—	4
160.0	5.9	—	3	182.0	2.0	—	3
160.0	5.3	—	3	182.0	1.6	—	4
160.0	4.86	+ 3.94 CO ₂	3	182.0	1.1	—	4
160.0	4.84	+ 2.5 CH ₃ CHO	3	182.0	1.05	—	3
160.0	4.66	+ 1.9 CH ₃ CHO	3	184.0	0.32	—	4*
160.0	4.21	+ 2.02 CO ₂	3	184.5	0.336	—	4*
160.0	4.20	+ 3.3 He	3	184.5	2.2	+ 3.8 He	3
160.0	4.13	—	3	184.5	2.08	+ 2.0 H ₂	3
160.0	4.13	—	3	184.5	1.7	—	3
160.0	4.04	—	3	184.5	1.6	—	3
160.0	3.64	+ 3.31 CO ₂	3	189	0.55	—	4*
160.0	3.59	+ 1.1 Air	3	189	0.195	—	4*
160.0	3.5	—	3				

TABLE II

Temp. ° C	k (min. ⁻¹)	Temp. ° C	k (min. ⁻¹)
184.5	0.430	156.5	0.0505
182	0.373	147.8	0.0216
176.5	0.227	144.4	0.0154
169.0	0.150	141.2	0.0116
160.0	0.065	140.2	0.0105

140° C). The reaction rate in the early stage of the reaction has been measured in most cases, because the rate tended to fall off as the peroxide was used up; this was found to be partly due to the mercury gauge method of pressure measurement; with the spoon gauge it was not noticed to nearly the same extent. (There was a very slight rise of temperature noticed on the thermometer inside the vessel after the commencement of the reaction indicating slight self heating; this also tends slightly to raise the reaction rate soon after the reaction has started.)

Very little difference was found in the results obtained for the slow reaction in the two vessels (see fig. 3), although the surface volume ratio of the one is 2.4 times that of the other. When the larger quartz vessel was one-third filled with small pieces of quartz, the reaction rate was unchanged. Also when the smaller pyrex vessel was coated with NaCl, the reaction rate remained the same as when it was uncoated. It may be concluded therefore that the decomposition occurs predominantly in the gas phase. Addition of helium up to 65 % of the total gas had no effect on the rate, neither had hydrogen (63 %) nor carbon dioxide (48 %) any effect. These results also indicate that the reaction is homogeneous and unimolecular and apparently uninfluenced by any chain process. Nevertheless, additions of nitric oxide increased the induction period which although short is quite definite, for instance at 169° C, 15 % NO increased the induction period from about 6 to 30 sec. and 30 % NO increased it to 2 min. Air appears to slow down the decomposition to a greater extent than can be accounted for by decrease in the pressure rise due to possible combustion. These effects may be due to the combination of radicals with NO or O₂. The influence of acetaldehyde on the reaction was also to slow down the reaction rate slightly, but its influence was erratic. Ethyl hydrogen peroxide on the other hand increased it considerably.

Experiments have been made at initial pressures as low as 2 mm. at 189° C without any falling off in the rate being observed. Special experiments therefore will be needed to ascertain the pressure at which the

decomposition of the peroxide ceases to behave according to a unimolecular law. Estimating from

$$N = \frac{Z \left(\frac{E + (\frac{1}{2}n - 1) RT}{RT} \right)^{(\frac{1}{2}n - 1)} \exp \left\{ - \left[\frac{E + (\frac{1}{2}n - 1) RT}{RT} \right] \right\}}{(\frac{1}{2}n - 1)!},$$

where N is the number of molecules reacting per sec. per c.c., Z is the number of collisions per sec. per c.c., E is activation energy, n is the number of squared terms for the peroxide, and, taking $n = 30$, the molecular diameter about 7×10^{-8} cm., and K at 190°C 0.67, the rate might be expected to fall off at about 0.17 mm. initial pressure. There was a slight falling off in the rate at 189°C as the partial pressure became less than 1.0 mm. It is noteworthy that azoisopropane showed no falling off in the value of the velocity constant even at 0.25 mm. initial pressure (Ramsperger 1928).

Although the decomposition of diethylperoxide appears to follow a unimolecular law down to the lowest pressures at which the process has been measured, yet like azomethane, above a certain critical pressure the reaction suddenly changes over to an explosion. The products are also different when the character of the reaction changes in this manner. The decomposition which occurs when the diethylperoxide is passed through a tube (1.1 cm. diameter) along with carbon dioxide leads to the formation of the products in Table III.

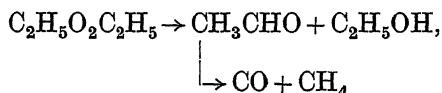
TABLE III

Temp. ° C	Et ₂ O ₂ reacted	HCHO	CH ₃ CHO	C ₂ H ₅ OH	CO	H ₂	CH ₄	C ₂ H ₆
187	8.65	1.9	3.2	7.5	4.7	Trace	2.7	2.3
201	17.4	3.9	6.9	14.2	9.8	1.1	5.9	3.9
245	28.4	16.8	7.9	10.4	6.5	0.4	7.8	7.4

The first two experiments agree quite well; in the first only 42 % of the peroxide reacted in the time of passage through the tube (11 cm. long), in the second experiment at a slightly higher temperature 56 % reacted. In the third experiment 100 % of the peroxide reacted and gave products in quite different proportions, corresponding with reaction above the critical limit. In this experiment the sum of the weight of the liquid products obtained from the analysis was 0.1178 g., while the total weight of the condensate was 0.1185 g. The ratio of the carbon, hydrogen and oxygen of the products is in all cases as 2 : 5 : 1, corresponding to the ratio in the original peroxide. The end gaseous products of the slow decomposition at 143°C are as follows: CO₂ 4.7 %, CO 55.6 %, CH₄ 19.4 %, C₂H₆ 19.4 %,

and not more than a trace of oxygen; so the quantity of CO_2 formed is less than one-tenth the amount of CO and even at this low temperature gaseous as well as condensable products are formed. Unsaturated compounds were not found in any of the above experiments.

It seems clear from these experiments that below the critical pressure the peroxide decomposes mainly as follows:

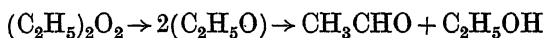


but that above the critical pressure, mainly as

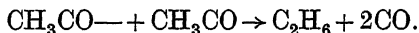
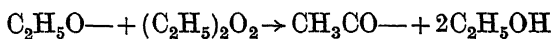


The explosive decomposition above the critical limit gives a final pressure about 2.8 times the critical pressure, the slow decomposition as already mentioned about 2.17 times the initial pressure.

The only figure in the above experiments which does not appear to be approximately accounted for by these reactions is the presence of rather more ethane in proportion to formaldehyde at the lower temperatures, possibly owing to decomposition of a proportion of the formaldehyde. It might be formed however from $(\text{CH}_3\text{CO})_2 \rightarrow \text{C}_2\text{H}_6 + 2\text{CO}$ (cf. Barak and Style (1935), who state that diacetyl was found in the products of the photochemical decomposition). Thus:



and



The critical explosion limits have been measured over a fairly wide range of temperatures and pressures in two quartz vessels. One vessel with flat ends was the same as used in the former experiments with vapour bath jacket, the other was a long quartz tube (1.4×45 cm.) in an electrically wound furnace. The flash was observed in the one case through one of the flat ends (fig. 1) and in the other through a long slit in the furnace jacket (fig. 6). The temperature was measured in the one case on a thermometer inside the vessel, in the other outside. By repeated trials, the maximum pressure at which the flash was not observed and the minimum pressure at which the flash was observed was determined, the actual pressure lying within the limits so obtained (see fig. 5).

The flash is light bluish in tint and at the higher pressures (above 2 cm.) a click is generally audible. Explosion appears to occur throughout the

vessel at the same moment. Very little lag was noticed on any occasion when the flash was obtained, never more than a second after letting in the vapour, and such lag as occurs is no doubt due to the time needed to reach the required explosion pressure. There was no explosion, only comparatively slow reaction at any pressure less than the critical pressure, however long the vapour was left in the vessel.

It has been possible to measure the slow reaction at temperatures well above the minimum temperature at which explosion has been observed, provided the initial pressure is less than the critical pressure for explosion. The critical pressures have been determined down to 176° C and the slow reaction measured up to 189° C. When explosion occurs reaction goes to completion, no further slow reaction is observed.

The results are given in Table IV.

TABLE IV

Vessel 1 (45 × 1.4 cm.)			Vessel 2 (6.5 × 4.5 cm.)		
Temp. ° C	Min. press. for explosion cm.	Max. press. for no explosion cm.	Temp. ° C	Min. press. for explosion cm.	Max. press. for no explosion cm.
193	4.10	4.02	176.5	3.65	3.25
197	3.62	3.50	181.4	2.40	2.18
204	3.50	2.45	184.5	2.00	1.90
211	2.02	1.95	185.0	2.04	1.94
218	1.75	1.70	189.0	1.71	1.63
225	1.30	1.22	191	1.75	1.70
239	1.0	0.97	193	1.43	1.40
253	0.62	0.57	200	0.80	0.75
267	0.4	0.35	202	0.83	0.83
290	0.3	(0.2)	207	0.64	0.60
			208.6	0.60	0.59
			210	0.54	0.49
			217.5	0.46	0.43
			223	0.32	0.31
			227	0.25	0.24
			234	0.21	0.18
			244	0.14	0.11

The experimental points have been plotted (fig. 5); the curves drawn through the points are the exponentials which refer to the straight lines:

$$\left. \begin{aligned} \log_{10} p &= \frac{5142}{T} - 10.912 \\ \text{or} \quad \log_e p &= \frac{23525}{RT} - 25.13 \end{aligned} \right\} \text{for vessel 2,}$$

$$\left. \begin{array}{l} \text{and} \quad \log_{10} p = \frac{3309}{T} - 6.503 \\ \text{or} \quad \log_e p = \frac{15138}{RT} - 14.974 \end{array} \right\} \text{for vessel 1,}$$

in which p is the pressure in cm. of mercury. The experimental points lie close about these curves, agreement being particularly good for the long vessel.

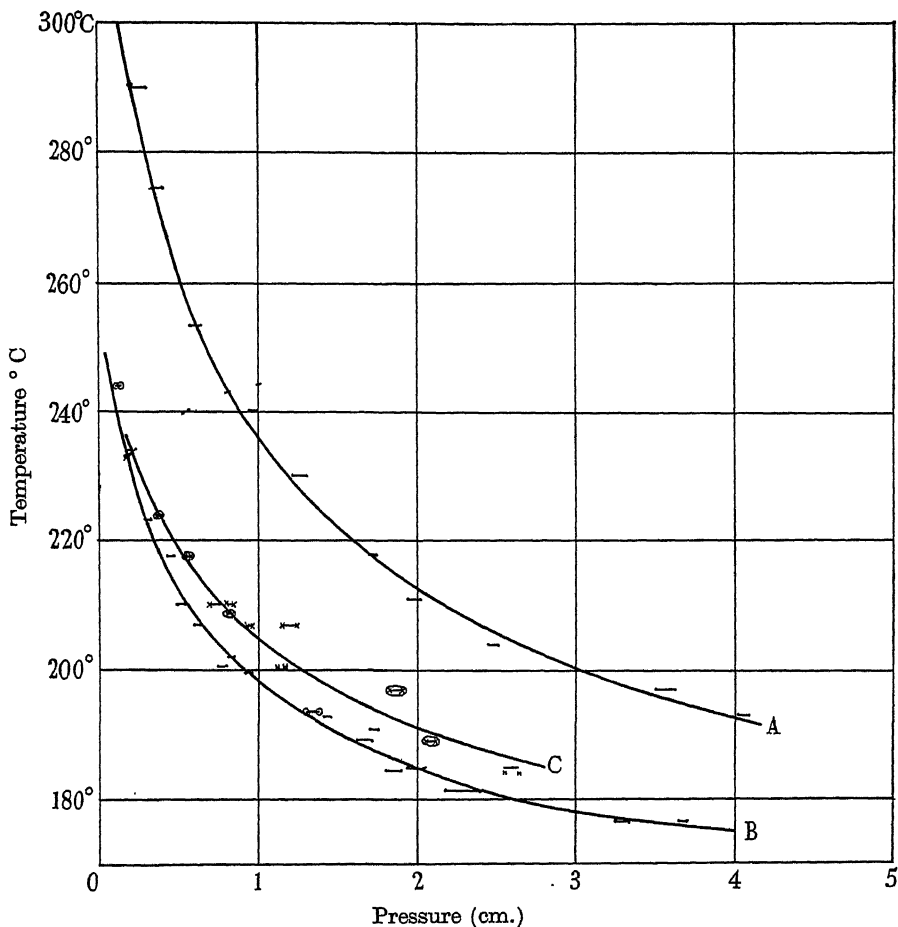


FIG. 5. Critical explosion pressures of diethyl peroxide. A, 1.4 cm. diameter vessel. B, 4.5 cm. diameter vessel. C, 50% He. HH hydrogen; x-x helium; (x-x) 50% He; o-o 50% CO₂; ••• (C₂H₅)₂O₂ alone.

Similar experiments have been made in the vessel of larger diameter on mixtures of diethylperoxide diluted with helium and other gases (Table V).

TABLE V

Temp. ° C	% diluent helium	Pressure (cm.)		Temp. ° C	% diluent	Pressure (cm.)	
		Min.	Max.			Min.	Max.
189	50	2.12	2.06	185	21 H ₂	2.27	2.27
193	22	1.41	1.37	185	44 H ₂	2.42	2.36
193	42	1.40	1.38	185	53 H ₂	2.66	2.56
193	65	—	1.5	200	33 H ₂	1.18	1.16
197	48	1.91	1.82	200	48 H ₂	—	1.26
207	66	0.96	0.94	193	33 CO ₂	1.37	1.31
207	86.5	1.26	1.15	193	39 CO ₂	1.39	1.29
208.6	55	0.84	0.82	244	84 CO ₂	0.13	0.11
210	57.5	0.7	0.7				
210	70	0.85	0.81				
217.5	50	0.58	0.56				
217.5	58.6	0.62	0.58				
217.5	72.5	0.75	0.70				
224	66	0.38	0.37				

The pressures given for critical explosion are those corresponding to the partial pressure of the diethylperoxide, the percentage of diluent is the percentage volume of diluent in the total gas.

The critical pressure is slightly affected when more than about 30 % of helium or of hydrogen are added, but with carbon dioxide no effect is found even up to 84 %. The graph (fig. 5) is drawn representing a mean through the figures for 50 % helium mixtures and corresponds to the equation

$$\log_e p = \frac{23174}{RT} - 24.40,$$

the slope therefore not being changed appreciably.

These results are somewhat similar to those obtained for azomethane (Allen and Rice 1935), the change-over from unimolecular reaction to explosion is explained in that case in terms of the Semenov theory of thermal explosions. They differ from the present results in that the rate does not appear to fall off at lower pressures as in the case of azomethane and no correction is needed to give the straight line agreement between the logarithm of the critical pressure and the reciprocal of the temperature, the induction lag is also shorter; in fact diethylperoxide is a particularly good example of the phenomenon where explosion suddenly at a certain pressure succeeds unimolecular decomposition.

The results are similar to those with azomethane as regards the effect of diluent gases and the effect of the surface/volume ratio which in both cases appears to have a larger influence than might be expected on existing theory. For instance, 50 % of helium which increases the conductivity

6 times raises the critical temperature for a particular pressure in both cases about 7 or 8°, while for an increase of surface/volume ratio 2.3 times the increase is up to 40° (for azomethane surface/volume ratio = 1.6 and rise in temperature about 20°). If the heat gained per sec. from the reaction per unit volume is $Qn(Ae^{-E/R(\Delta T+T)})$, where ΔT is the rise in temperature above T the temperature of the walls, E is the activation energy, Q the heat of dissociation per mol, and n the number of molecules per unit volume, and if the heat lost per sec. per unit volume is $(a/v) b \Delta T$, where a/v is the surface volume ratio and b a constant, and if the critical explosion pressure depends on the balance of these quantities, then the above observations about the effect of tube diameter are not accounted. Nevertheless the effect of the diluents, the homogeneous nature of the slow reaction, and the exothermic character of the decomposition make it probable that the explosion can be explained thermally without the introduction of any radical chain mechanism.

The thermal theory outlined is unsatisfactory because it does not take into account (1) that the outer layer of gas can exchange heat more readily with the walls than inner layers, (2) that the gas has to be heated through the agency of collision with the walls, (3) that the reactant may not be able to lose activation energy to the products or diluents, (4) that there may be alternative modes of decomposition requiring a different activation energy, (5) that the conductivity may be influenced by movements in the gas and (6) that part of the energy may be removed by radiation.

With a view to ascertaining more precise knowledge of the conditions affecting this type of explosion, it is proposed to make more careful experiments on the effect of the diameter of the explosion vessel on the critical pressure. The effect of increase of the percentage of diluents is also being further investigated, for the experiments indicate that even at great dilutions provided a critical partial pressure is attained the slow will give place to the explosive reaction. The formula would indicate that concentrations of the order 10^{-5} might decompose explosively below 400° C.

Ignition of diethylperoxide

Experiments have already been recorded (Table I) in which the effect of air on the slow decomposition of diethylperoxide was tested: the decomposition became slower apparently, as the following figures show for an experiment at 157° C:

Time	With air	Without air
	$\log \frac{a}{a-x}$	$\log \frac{a}{a-x}$
34 sec.	0.008	0.028
2 min. 13 sec.	0.049	0.11

In air at higher temperatures, diethylperoxide ignites and flame is propagated. Ignitions were carried out in a quartz tube (45×1.4 cm.) (fig. 6), heated in an electrically wound heating jacket furnished with a longitudinal slit. The peroxide-air mixture at a measured pressure is let into the evacuated tube. Ignition starts at the far end away from the inlet tap (where the tube was slightly hotter) and a bluish flame travels back through the gas mixture. The ignition phenomenon appears quite different from the flash throughout the gas which is associated with the critical pressure decomposition phenomenon. The following figures apply to mixtures of diethylperoxide in about 10 cm. total pressure of air:

TABLE VI

% of peroxide	Ignition temp. ° C
11	200
3.5	230
1.6	260
0.9	280

In the last experiment the peroxide is igniting at a partial pressure less than 1 mm.

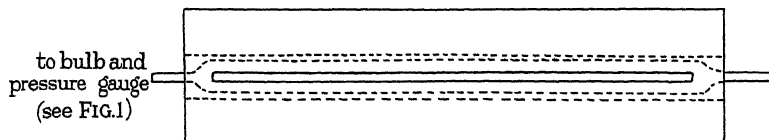


FIG. 6

Similar experiments have been made using a 1 : 1 propane oxygen mixture instead of air in which to burn the peroxide:

TABLE VII. PRESSURE OF PROPANE OXYGEN MIXTURE 10 CM. APPROX.

Pressure of peroxide (cm.)	Ignition temp. ° C
2.07-2.25	190
1.35-1.45	200
0.9 -1.02	211
0.5 -0.6	232
(0.35-0.4)	(260)
0.10-0.2	290
0.0 -0.05	325

The temperature of ignition of the peroxide is raised as a result of the addition of propane instead of nitrogen, nevertheless ignition continues to very low partial pressures of peroxide.

At the lowest pressures a blue glow is visible at the entry to the tube, a flash occurs but no further flame combustion. Under these conditions the propane itself does not ignite and even at the higher temperatures the propane does not show any sign of oxidation until several minutes after the ignition of the peroxide. The peroxide flame in fact does not ignite the propane at these temperatures.

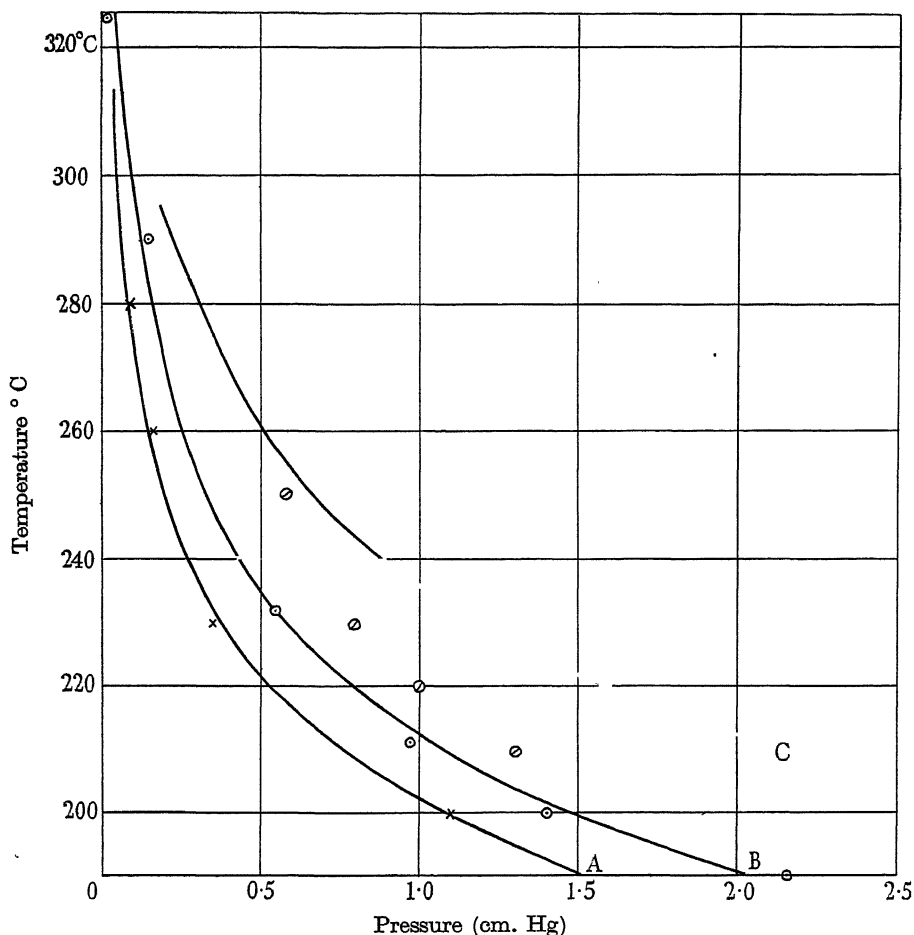


FIG. 7. Ignition of diethyl peroxide. A, In air. B, In propane oxygen 1:1. C, Critical explosion (alone). o, Neumann and Toutakin.

The results of the above experiments are plotted in fig. 7, also the graph for the explosive decomposition of diethylperoxide in the same size of tube is reproduced (see also fig. 5). The graphs drawn correspond to the following equations connecting $\log p$ and $1/T$:

(a) For diethylperoxide alone:

$$\log_e p = \frac{15140}{RT} - 14.97.$$

(b) For the diethylperoxide in presence of 1 : 1 propane oxygen:

$$\log_e p = \frac{14850}{RT} - 15.4.$$

(c) For the diethylperoxide in presence of air:

$$\log_e p = \frac{16300}{RT} - 17.3.$$

Clearly the critical pressure phenomenon is controlling the ignition, as the slopes are all very similar. In air ignition occurs at about one half the pressure of diethylperoxide as is required to obtain explosive decomposition, the energy of combustion being available to favour thermal ignition; in the propane mixture, however, the ignition does not occur quite so readily as in air. The flame is very similar in appearance. Neumann and Toutakin's (1937) results for the critical explosion pressure of diethylperoxide are given

by the relation $\log_e p = \frac{10500}{RT} - 9.5$; the slope of the graph is very different from those obtained in this investigation, possibly owing to a different shape of vessel (the diameter was stated to be 3.5 cm.).

They found that ignition and cool flame was engendered by diethylperoxide in a 1 : 1 butane oxygen mixture at 30 cm. pressure and the pressures of diethylperoxide then corresponded to the pressure for critical explosion. Although it would no doubt be possible to ignite propane mixtures through the decomposition of diethylperoxide at lower temperatures than for a similar propane mixture alone, as is indicated by the effect of diethylperoxide on the induction period, yet, as here shown, it is possible to obtain explosive decomposition of the diethylperoxide without ignition of the propane.

These observations provide the basis of an explanation of the many apparently complicated and curious phenomena associated with the production and propagation of cool flames. If products are formed in a combustible mixture by partial oxidation which behave similarly to diethylperoxide, they could propagate flame in the mixtures provided their partial pressures were above certain values though the actual partial pressures might be quite small. Flames might follow each other as more of the substance was formed or if sufficiently intense they would ignite the rest of the gas and give rise to the ignition of the mixture.

The peroxide greatly diminishes the induction period of the propane oxygen (1 : 1 mixture) slow combustion which occurs at higher temperatures:

TABLE VIII. TEMP. 350° C. QUARTZ VESSEL

Pressure of propane and oxygen (cm.)	Pressure of peroxide cm.	Percentage	Induction time sec.
39.1	Nil	—	220
39.5	Nil	—	240
38.5	Nil	—	285
38.8	0.06	0.15	30
38.2	0.11	0.29	< 10
39.0	0.11	0.28	8
40.8	0.10	0.27	< 10

The peroxide or its combustion products therefore increase the number of reaction centres for the propane oxidation: 0.25 % of diethylperoxide diminishes the induction period to 1/30th its previous value.

Further experiments are being made to extend these investigations and follow up some of the indications here mentioned. The behaviour of other peroxides such as ethylhydrogen peroxide and peracetic acid is also being investigated.

SUMMARY

1. The decomposition of diethylperoxide, which measured in the temperature range 130–190° C, behaves unimolecularly, appears to be unaffected by surface or diluents. The apparent activation energy is 31,500 cal. and $\log_{10} K = -\frac{6890.4}{T} + 14.708$.

2. The decomposition becomes explosive at a certain critical pressure, the dependence on temperature of which is given by

$$\log_{10} p \text{ cm.} = \frac{5142}{T} - 10.912$$

in one of the quartz vessels used. The critical explosion pressure is only slightly raised by conducting diluents such as hydrogen or helium but is affected greatly by the diameter of the vessel.

3. Diethylperoxide ignites in air and gives rise to flame at temperatures somewhat lower than correspond to the critical explosion pressures in a neutral atmosphere in the same vessel.

4. Ignition occurs in propane oxygen mixtures at temperatures below those at which the propane itself will ignite, without igniting the propane.

5. Diethylperoxide greatly diminishes the induction period of the slow combustion of propane.

REFERENCES

- Allen, A. O. and Rice, O. K. 1935 *J. Amer. Chem. Soc.* **57**, 310.
Baeyer and Villiger 1900 *Ber. dtsh. chem. Ges. B*, **33**, 3387.
Barak, M. and Style, D. W. G. 1935 *Nature, Lond.*, **135**, 307.
Callendar, H. L. 1927 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1062.
Egerton, A. C. 1927 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1079.
— 1928 *Nature, Lond.*, Suppl. **122**, 20.
Egerton, A. C., Smith, F. L. and Ubbelohde, A. R. 1935 *Philos. Trans. A*, **234**, 433.
Harris, E. J. and Egerton, A. C. 1937 *Chem. Rev.* **21**, 287.
Medvedeff, S. and Podjapolskaya, A. 1935 *Acta Phys. Chem. U.R.S.S.* **2**, 487.
Mondain Monval, P. and Quanquin, B. 1929 *C.R. Acad. Sci., Paris*, **189**, 1194.
— — 1931 *Ann. Chemie*, **15**, 309.
Neumann, M. and Aivazov, B. 1935 *Nature, Lond.*, **135**, 655.
Neumann, M. and Toutakin 1937 *C.R. Acad. Sci., Paris*, **205**, 278.
Pease, R. N. 1934 *J. Amer. Chem. Soc.* **56**, 2034.
— 1935 *J. Amer. Chem. Soc.* **57**, 2296.
Ramsperger 1928 *J. Amer. Chem. Soc.* **50**, 714.
Rieche, A. 1931 "Alkylperoxyde und Ozonide". Steinkopff.
Townend, D. T. A. 1933 *Proc. Roy. Soc. A*, **143**, 168.
— 1938 *J. chem. Soc.* 238.
-

Formation of negative ions at metal surfaces

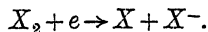
By R. A. SMITH, PH.D.

Carnegie Teaching Fellow, St Andrews University

(Communicated by N. F. Mott, F.R.S.—Received 10 March 1938)

In a previous paper (Massey and Smith 1936) a theoretical treatment of various processes of formation of negative ions in gases has been given. The recent experimental work of Arnot (1936, 1937) has, however, shown the importance of formation taking place at the surface of metal electrodes. The purpose of this paper is to give a theoretical treatment of the formation of negative ions at metal surfaces and to indicate where further experimental research is necessary in order to understand more fully the processes taking place.

When a neutral atom captures an electron to form a negative ion, excess energy equal to the sum of the initial kinetic energy of the electron and the electron affinity of the atom must be disposed of. The method of disposal largely determines the probability with which a process will take place. These processes may be divided into two classes, radiative and non-radiative. In the former the excess energy is given off as radiation. The effective cross-section for radiative attachment to atoms is of the order of 10^{-22} cm.² for electrons having energies of a few electron volts (Massey and Smith 1936). When negative ions are formed in a monatomic vapour such as mercury, the only other means of disposal of the excess energy (excluding surface effects) is through a three-body collision involving another atom or electron. This has been shown to have a very small chance of taking place under the conditions usually met with in electrical discharges (Smith 1936). In polyatomic gases, however, non-radiative processes may result from a single collision. We should expect these to take place with considerably greater probability than those involving radiation. The most common is that of formation of atomic negative ions by molecular dissociation in reactions of the type



The excess energy in this process goes into the kinetic energy of the products of dissociation. A qualitative treatment of such reactions has been given by Massey and Smith (1936). They account for most of the experimental results on the formation of atomic negative ions by direct electron impact

in polyatomic gases. When molecular negative ions are formed in these gases by attachment of free electrons, without dissociation, the process must be either that of radiative capture or that suggested by Bloch and Bradbury (1935), whereby the excess energy is taken up in exciting the molecule to higher vibrational states and is then lost by collision with other molecules.

Arnot and Milligan (1936) have shown that the above processes cannot account for the negative ions formed in mercury under the conditions of their experiments, and that a process which takes place with considerably higher probability is required. This process was shown to be the capture of two electrons by positive ions, when they strike the surface of a metal electrode. Arnot has observed this for several polyatomic gases as well as for mercury (1937). We shall see that this is a resonance process, there being no excess energy to dispose of. Accordingly it will take place with a comparatively large probability under suitable conditions. Arnot found that the chance of a positive ion being converted into a negative ion varies between about 10^{-3} and 10^{-5} for different gases.

In recent papers (see, for example, Emeleus and Sayers 1938) it has been shown that negative ions play a large part in determining the form of certain types of electrical discharges. Clearly the processes taking place at the electrodes are of great importance as, according to Arnot's experiments, they account for a considerable number of the negative ions in the discharge.

1. THE PROCESSES OF FORMATION AND NEUTRALIZATION OF NEGATIVE IONS AT METAL SURFACES

When a positive ion strikes a metal surface, in general it is neutralized by capturing an electron from the metal (Massey 1930). The reverse process also may take place. When a caesium atom strikes a hot tungsten surface, it loses an electron and is given off as a positive ion. The conditions for the occurrence of these processes are respectively $V_i > \phi$ and $V_i < \phi$, where V_i is the ionization potential of the neutral atom or molecule which is formed, and ϕ the work function of the metal. It is known that when an ion such as He^+ is neutralized by the above process the electron may be captured into an excited state, and on falling into the ground level may eject an electron from the metal. This process has been examined experimentally by Oliphant (1929, 1930), and a quantum-theoretical treatment has been given by Massey (1930, 1931). Instead of being ejected from the metal, under certain conditions, the second electron may be captured to form a negative ion, the energy of excitation being used to extract the electron from the metal as

before. The process will clearly take place with appreciable probability only when there is an occupied level in the metal from which an electron may be captured so that it has just the right energy to form a negative ion. This is undoubtedly the process taking place in Arnot and Milligan's experiments on the formation of mercury negative ions. We shall now proceed to discuss other possible processes of negative-ion formation at metal surfaces and to specify the conditions under which they may take place.*

(1) *Formation from neutral unexcited atoms.* If there is an occupied level in the metal from which an electron may be captured by resonance to form a negative ion, this process will have a large chance of taking place when the atom comes near to the surface of the metal. This is exactly analogous to the neutralization of a positive ion. The condition which must be satisfied is

$$\mu + \phi > W_e > \phi, \quad (a)$$

μ being the maximum kinetic energy of a "free" electron in the metal (upper limit of Fermi distribution) and ϕ the work function as before. W_e is the electron affinity of the atom (see fig. 1). This process will not be of much importance in practice. The only case in which (a) is likely to be satisfied

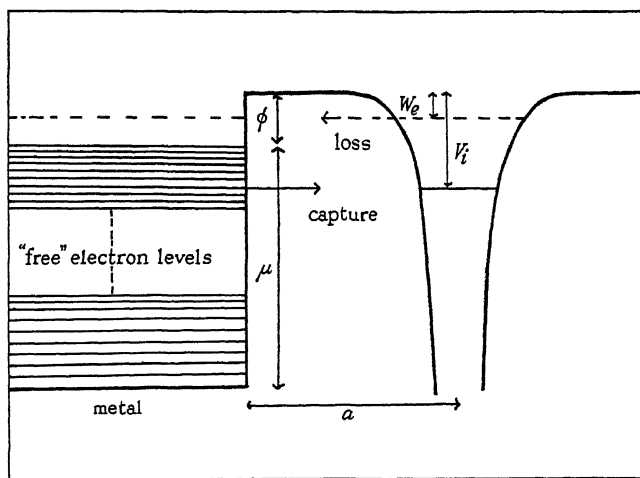


FIG. 1. Illustrating neutralization of a positive ion and a negative ion near the surface of a metal.

* [Note added in proof. These processes are supposed to take place at a clean metal surface. When a good deal of adsorbed gas is present more complicated processes take place involving transfer of energy from the incident positive ions to the adsorbed gas atoms, some of which may be ejected as negative ions. (Arnot and Beckett 1938, *Nature*, **141**, 1011.)]

will be that of halogen atoms incident on a clean alkali metal surface. This, however, would not remain long uncontaminated in the presence of halogen atoms.

(2) *Formation from excited atoms.* When the incident atom is excited, the energy of excitation is available to extract the electron from the metal, and the condition to be satisfied is

$$\mu + \phi > V_e + W_e > \phi, \quad (b)$$

where V_e is the excitation potential of the state involved. Clearly the condition (b) will be much more frequently satisfied than (a). This process undoubtedly takes place in discharges, although there is little experimental information about it. Metastable states will of course be most effective. There is a considerable probability, as we shall see, of slow negative ions being neutralized near a metal surface. Only excited atoms which have considerable kinetic energy will therefore be effective. This may account for the fact that the process has not been observed with certainty in discharges.

(3) *Formation from positive ions which first form excited neutral atoms.* This process has been discussed above.

The condition for the formation of an excited atom from a positive ion is

$$\mu + \phi > V_i - V_e > \phi. \quad (c)$$

The condition for the subsequent formation of the negative ion is given by (b). Together these give

$$2\mu + 2\phi > V_i + W_e > 2\phi. \quad (d)$$

We must notice, however, that (d) may be satisfied when (c) is not. This is so for atomic hydrogen and nickel. For the first excited state $V_e = 3.37$ eV, $V_i = 13.5$ eV, $W_e = 0.7$ eV and for Ni $\phi = 5.03$ eV (Fox and Bowie 1933). We see that (b) and (d) are satisfied but that (c) is not. For mercury both (c) and (d) are satisfied for the 2^3P_0 metastable state, the values of V_e and V_i being 4.66 and 10.38 eV respectively.

(4) *Formation from positive ions by simultaneous capture of two electrons.* It is possible for a positive ion to capture two electrons from the metal in a single transition without going through the intermediate stage of forming an excited atom. The condition for this is simply given by (d). This process is therefore sometimes energetically possible when (3) is not. We shall give reasons, however, for expecting its probability of occurrence to be small.

(5) *Formation from molecular ions.* When the ions striking the metal surface are polyatomic, the processes taking place are more complicated as dissociation may take place. Arnot found that both O^- and O_2^- ions were

formed from O_2^+ , and O^- , O_2^- , CO^- and C^- from CO_2^+ . When the negative ions are formed without dissociation, the conditions (a)–(d) simply apply to the incident molecules or ions, but if these break up they must be applied to the products of dissociation.

(6) *Neutralization of negative ions.* After a negative ion has been formed near a metal surface it may be neutralized by a process which is the converse of (1), i.e. the extra electron of the negative ion may pass into an unoccupied level of the same energy in the metal. This will be possible whenever (1) cannot take place, the condition being

$$W_e < \phi \quad (e)$$

(see fig. 1). This process is the analogue of the ionization of caesium on tungsten. A negative ion striking an ordinary metal surface will thus in general be neutralized. Moreover, a considerable number of the negative ions formed by processes (2), (3) and (4) will be lost before they escape from the surface.

2. FORMATION OF MERCURY NEGATIVE IONS AT A NICKEL SURFACE*

We choose the formation of mercury negative ions at a nickel surface for detailed calculation, since this is the only monatomic vapour for which accurate experimental information is available. Clearly process (1) cannot take place as the electron affinity as calculated by Glocker (1934) is 1.9 eV, and this value is probably too high. Arnot and Milligan have shown that if process (2) takes place it is not appreciable compared with (3) or (4) under the conditions of their experiments. The latter are not experimentally distinguishable in mercury, as (b), (c), (d) are all satisfied. We shall first consider process (3).

The method of calculation we shall adopt is an extension of that used by Massey (1930, 1931) in his treatment of ejection of electrons from metals by metastable atoms. For the metal we shall use the Sommerfeld model (see fig. 1). Since the incident positive ions are moving slowly compared with the electrons involved in the various transitions for which we shall require to calculate the probability per unit time, we shall first evaluate these assuming the ion or excited atom to be at rest at distance a from the surface of the metal.

* [Note added in proof. The results for tungsten are almost identical. Tungsten is more suitable for experiment as it may be made clean more readily than nickel. (See footnote p. 21.)]

First we shall calculate the probability of a transition of a "free" electron in the metal to an excited state in the field of the incident positive ion. Let \mathbf{r} represent the co-ordinates of the metallic electron and \mathbf{r}_s the co-ordinates of the electrons of the positive ion referred to its centre of mass as origin. Let $\psi_0(\mathbf{r}, \mathbf{r}_s)$ be the wave function describing the "free" electron in the metal and the electronic state of the positive ion. Let $\psi_f(\mathbf{r}, \mathbf{r}_s)$ be the wave function for the excited neutral atom formed by the transition. Massey has shown (1930) that the required transition probability per unit time $p(a)$ is given by the expression

$$p(a) = \frac{6\pi^2 N \bar{v} E^{\frac{1}{2}}}{\hbar \mu^{\frac{3}{2}}} |V_{of}|^2, \quad (1)$$

where V_{of} is the matrix element of the interaction energy V under which the transition takes place, i.e.

$$V_{of} = \int \psi_0 V \psi_f d\mathbf{r} d\mathbf{r}_s. \quad (2)$$

N is the number of "free" electrons per unit volume, \bar{v} the average velocity, and μ the maximum kinetic energy of electrons in the metal. E is the kinetic energy of the electron involved in the transition and is given by the equation

$$\mu + \phi - E = V_i - V_e. \quad (3)$$

The expression (1) takes no account of electron exchange. This may be done by replacing V_{of} by

$$\frac{3}{4} |V_{of} - V'_{of}|^2 + \frac{1}{4} |V_{of} + V'_{of}|^2, \quad (4a)$$

when two electrons whose co-ordinates are \mathbf{r}_1 and \mathbf{r}_2 are considered. V'_{of} is obtained from V_{of} by interchanging \mathbf{r}_1 and \mathbf{r}_2 in ψ_f . For transitions involving three electrons (one free and two bound) V_{of} is replaced by

$$|V_{of} - V'_{of}|^2. \quad (4b)$$

(Oppenheimer 1928).

Next we shall require the transition probability $m(a)$ for the formation of a negative ion through the excited atom falling into its ground state and capturing another electron from the metal. This will be given by an expression exactly similar to (1), but $\psi_0(\mathbf{r}, \mathbf{r}_s)$ now refers to the free electron (\mathbf{r}) and the excited atom (\mathbf{r}_s), and $\psi_f(\mathbf{r}, \mathbf{r}_s)$ to the negative ion. E will now be given by the equation

$$\mu + \phi - E = V_e + W_e. \quad (5)$$

Lastly we must calculate the transition probability corresponding to neutralization of the negative ion by process (6). This again will have the form (1) with E given by the equation

$$\mu + \phi - E = W_e. \quad (6)$$

The reason for this is that (1) was obtained on the assumption that all possible states in the metal with energy near E are occupied, and in the reverse process all the corresponding states are available since $W_e < \phi$. ψ_0 now refers to the negative ion and ψ_f to a "free" electron in the metal and a normal atom.

We shall now consider the motion of the positive ion. For our purpose it will not be sufficient to integrate the transition probabilities over all distances a from the surface as in Massey's calculation on the neutralization of positive ions. We must set down the differential equations satisfied by the probabilities of the incident ion being in each of the states we consider, at distance a from the surface. We shall treat only the case of normal incidence. We shall do this in two parts: (a) when the incident ion is moving towards the surface, (b) when it (or one of its products) is moving away from the surface after impact. Let $P(a)$ be the probability that an incident positive ion remains unneutralized at distance a from the surface, $M(a)$ that of its being an excited atom, and $N(a)$ that of its being a negative ion. Then we have

$$\left. \begin{aligned} v_0 \frac{dP}{da} &= p(a) P, \\ v_0 \frac{dM}{da} &= m(a) M - p(a) P, \\ v_0 \frac{dN}{da} &= n(a) N - m(a) M, \end{aligned} \right\} \quad (7)$$

where v_0 is the velocity of the ion towards the surface. In forming these equations we have neglected the chance that the positive ion will be neutralized by capturing an electron to form a normal atom. Massey has shown, however, that the most likely state is the highest permitted by the energy condition (c). We have also neglected the chance that the excited atom will eject an electron from the metal on falling into its ground state. As the available energy is only 0.69 eV for the 2^3P_0 metastable state we should expect the chance of this happening to be small. There is some experimental evidence that this is so (Compton and Langmuir 1930).

When a positive ion or one of its products reaches the surface it normally loses most of its kinetic energy. The experiments of Arnot show, however, that some are reflected with a considerable amount of their initial energy. The collision processes taking place will be very complicated, and we must make some simplifying assumptions. We shall assume that a particle moves up to the surface with constant velocity v_0 and that it can be reflected with velocity between 0 and v_0 . We must assume some distribution function to

give the chance of reflexion with velocity between v and $v + dv$. This model is admittedly crude as the surface cannot be regarded as plane when a is of the order of 10^{-8} cm. Also the velocity of the incident particles will undergo rapid changes in magnitude and direction as they collide with the surface atoms. However, we may expect it to give a fair representation of the average effect for a large number of incident ions.

With these assumptions we now form equations for the outgoing particles similar to (7). As we shall see, most of the positive ions are neutralized before reaching the surface, so we can now neglect $P(a)$. Let $M'(a)$, $N'(a)$ be the probabilities of the incident positive ion being an excited atom and negative ion at distance a from the surface, after reflexion with velocity v . For $M'(a)$, $N(a)$ we have the equations

$$\left. \begin{aligned} v \frac{dM'}{da} &= -m(a) M', \\ v \frac{dN'}{da} &= -n(a) N + m(a) M'. \end{aligned} \right\} \quad (8)$$

The appropriate solutions of the equations (7) are as follows

$$P(a) = 1 - \exp \left\{ -\frac{1}{v_0} \int_a^\infty p(a) da \right\}, \quad (9)$$

$$M(a) = \exp \left\{ -\frac{1}{v_0} \int_a^\infty m(a) da \right\} \int_a^\infty p(a) P(a) \exp \left[\frac{1}{v_0} \int_a^\infty m(x) dx \right] da, \quad (10)$$

$$N(a) = \exp \left\{ -\frac{1}{v_0} \int_a^\infty n(a) da \right\} \int_a^\infty m(a) M(a) \exp \left[\frac{1}{v_0} \int_a^\infty n(x) dx \right] da. \quad (11)$$

Let P_0 , M_0 , N_0 be the solutions when $a = 0$. These give the chance of an incident ion reaching the surface as a positive ion, excited atom, and negative ion respectively. Then the appropriate solutions of (8) are

$$M'(a) = M_0 \exp \left\{ -\frac{1}{v} \int_0^a m(a) da \right\}, \quad (12)$$

$$\begin{aligned} N'(a) &= N_0 \exp \left\{ -\frac{1}{v} \int_0^a n(a) da \right\} \\ &+ \exp \left\{ -\frac{1}{v} \int_0^a n(a) da \right\} \int_0^a m(a) M'(a) \exp \left[\frac{1}{v} \int_0^a n(x) dx \right] da. \end{aligned} \quad (13)$$

These solutions depend on the assumption that the state of a given particle is not changed by the actual collision with the surface. The probability of such electronic transitions during the impact of two heavy particles is known to be small when their kinetic energy of relative motion is of the order of 100 eV (Massey and Smith 1933), and we should expect the same to hold in the case of impact with the surface of the metal.

The values of $M'(a)$, $N'(a)$, when a is large, give the probabilities that the incident positive ion will leave the surface with velocity v as an excited atom and negative ion respectively. For a given value of v_0 these will be functions of v . It is more convenient to express them in terms of W , the kinetic energy of the reflected ions. Let $N(V, W)$ be the chance that a positive ion whose kinetic energy was V before collision with the surface and W after, leaves the surface as a negative ion. Let $G(V, W)dW$ be the probability that a particle, incident with energy V , will come off with energy between W and $W + dW$. The energy distribution of *negative ions*, that is, the number escaping from the surface with energy between W and $W + dW$, when a large number of positive ions with energy V strike the surface, is given by $G(V, W) N(V, W) dW$. To obtain the average probability of formation $\bar{N}(V)$ we shall have to integrate over all values of W between 0 and V . We shall then have

$$\bar{N}(V) = \int_0^V G(V, W) N(V, W) dW. \quad (14)$$

$\bar{N}(V)$ and $G(V, W) N(V, W)$ are the quantities given by Arnot and Milligan's experiments. The function $G(V, W)$ is not given directly and we shall have to assume some form for it. For mercury we shall see that good agreement with experiment is obtained with the simple form

$$G(V, W) = Ce^{-W/W_0}, \quad (15)$$

where W_0 is a function of V and is the mean energy of the reflected particles, and C is such that

$$\int_0^V G(V, W) dW = 1.$$

Calculation of $p(a)$, $m(a)$, $n(a)$, for mercury

(1) $p(a)$. The wave functions which we take to describe the initial and final states are as follows:

$$\psi_0 = \chi(\mathbf{r}_2) \phi(\mathbf{r}_1),$$

where

$$\begin{aligned} \chi(r_2) &= u^{-\frac{1}{2}} e^{ik(\mathbf{a}, \mathbf{r}_2)}, & r > a \sec \theta, \theta < \frac{1}{2}\pi \\ &= 0, & \text{otherwise,} \end{aligned} \quad (16)$$

and
$$\phi(\mathbf{r}_1) = N_0 r_1^{t-1} e^{-z' r_1 / t}, \quad (17)$$

$k = 2\pi mu/h$, u being the velocity of the "free" electron and \mathbf{n} a unit vector along the direction of its initial motion. u is obtained by putting $E = \frac{1}{2}mu^2$ in equation (3),

$$\psi_f^{\alpha, \beta, \gamma}(\mathbf{r}_1, \mathbf{r}_2) = N_1 r_1^{s-1} r_2^{s-1} e^{-z(r_1+r_2)/s} \begin{cases} \cos \theta_1 - \cos \theta_2, \\ \sin \theta_1 e^{\pm i\phi_1} - \sin \theta_2 e^{\pm i\phi_2}. \end{cases} \quad (18)$$

The wave function $\chi(\mathbf{r}_2)$ represents the "free" electron in the metal and $\phi(\mathbf{r}_1)$ the mercury positive ion. $\psi_f^{\alpha, \beta, \gamma}(\mathbf{r}_1, \mathbf{r}_2)$ represent the 2^3P_0 state of mercury. It will be noticed that only the two electrons on the outer shell have been included. Those in deeper levels will not be much affected by the transition. The constants s, t, z, z' , are obtained by means of Slater's rules (Slater 1930). The use of wave functions of this type for calculations on negative ions has been discussed by Massey and Smith (1936). To perform the calculation with the accurate form (16) for $\chi(\mathbf{r}_2)$ would be extremely laborious, and following Massey (1930) we shall treat the case of an ion at rest at the centre of a spherical cavity in the metal, replacing the conditions for the vanishing of $\chi(\mathbf{r}_2)$ by

$$\chi(\mathbf{r}_2) = 0, \quad r_2 > a. \quad (19)$$

For the interaction energy we take

$$V(\mathbf{r}_1, \mathbf{r}_2) = \frac{2e^2}{r_2} - \frac{e^2}{r_{12}}, \quad (20)$$

making the approximation that the screening factor for electrons in the inner shells is unity. For the three degenerate states represented by $\psi_f^{\alpha, \beta, \gamma}$ we must replace $|V_{0f}|^2$ by

$$|V_{0f}^{\alpha}|^2 + |V_{0f}^{\beta}|^2 + |V_{0f}^{\gamma}|^2, \quad (21)$$

where $V_{0f}^{\alpha, \beta, \gamma}$ are formed with the functions $\psi_f^{\alpha, \beta, \gamma}$ respectively. We choose the direction of motion of the "free" electron as the axis $\theta_2 = 0$. It can then be shown that $V_{0f}^{\beta} = V_{0f}^{\gamma} = 0$. For V_{0f}^{α} we have the expression

$$V_{0f}^{\alpha} = \frac{N_0 N_1 e^2}{u^{\frac{1}{2}}} \int_{r_2 < a} d\mathbf{r}_2 \int_{r_1^{s+t-2} r_2^{s-1} \left(\frac{2}{r_2} - \frac{1}{r_{12}} \right) \exp \left\{ i k r_2 \cos \theta_2 - \left(\frac{z}{s} + \frac{z'}{t} \right) r_1 - \frac{z}{s} r_2 \right\} \times (\cos \theta_1 - \cos \theta_2) d\mathbf{r}_1. \quad (22)$$

To evaluate this integral we expand $V(\mathbf{r}_1, \mathbf{r}_2)$ in the series

$$e^2 \sum_{n=0}^{\infty} \sum_{m=0}^n \gamma_n(r_1, r_2) P_n^m(\cos \theta_1) P_n^m(\cos \theta_2) \cos m(\phi_1 - \phi_2) (l-m)!/(l+m)!, \quad (23)$$

where

$$\gamma_0 = \frac{1}{r_2}, \quad r_1 < r_2, \quad \gamma_n = -\frac{r_1^n}{r_2^{n+1}}, \quad r_1 < r_2, n > 0$$

$$= \frac{2}{r_2} - \frac{1}{r_1}, \quad r_1 > r_2, \quad = -\frac{r_2^n}{r_1^{n+1}}, \quad r_1 > r_2, n > 0.$$

Integrating with respect to θ_1 and ϕ_1 we have

$$V_{0f}^\alpha = \frac{4\pi^2 e^2 N_0 N_1}{u^{\frac{1}{2}}} \int_{r_2 < a} d\mathbf{r}_2 \int_0^\infty r_1^{s+t} r_2^{s-1} \cos \theta_2 \left(\frac{1}{3} \gamma_1 - \gamma_0 \right) \\ \times \exp \left\{ ikr_2 \cos \theta_2 - \left(\frac{z}{s} + \frac{z'}{t} \right) r_1 - \frac{z}{s} r_2 \right\} dr_1. \quad (24)$$

Using the expansion

$$e^{i kr \cos \theta} = \frac{\pi^{\frac{1}{2}}}{(2kr)^{\frac{1}{2}}} \sum_{n=0}^{\infty} (2n+1) i^n J_{n+\frac{1}{2}}(kr) P_n(\cos \theta), \quad (25)$$

we obtain

$$V_{0f}^\alpha = 16\pi^2 e^2 N_0 N_1 \left(\frac{\pi}{2ku} \right)^{\frac{1}{2}} \int_a^\infty dr_2 \int_0^\infty J_{\frac{3}{2}}(kr_2) r_2^{s+\frac{1}{2}} r_1^{s+t} \left(\frac{1}{3} \gamma_1 - \gamma_0 \right) \\ \times \exp \left\{ - \left(\frac{z}{s} + \frac{z'}{t} \right) r_1 - \frac{z}{s} r_2 \right\} dr_1. \quad (26)$$

The final integration is best carried out numerically.

The integrals in the calculation of V'_{0f} are exactly similar to those we have evaluated. We may then proceed to calculate $p(a)$ using the expression (4b).

(2) $m(a)$. Here

$$\psi_0 = \chi(\mathbf{r}_3) \psi(\mathbf{r}_1, \mathbf{r}_2), \quad (27)$$

where $\psi(\mathbf{r}_1, \mathbf{r}_2)$ is the function (18) and $\chi(\mathbf{r})$ the function (16) with u obtained from the value of E given by equation (5). The wave function which we take for the Hg^- ion is

$$\psi_{f,\beta,\gamma}^{\alpha,\beta,\gamma}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = N_2 r_1^{p-1} r_2^{p-1} r_3^{q-1} \exp \left[-\frac{\xi}{p} r_1 - \frac{\xi}{p} r_2 - \frac{\xi'}{p} r_3 \right] \begin{cases} \cos \theta_3, \\ \sin \theta_3 e^{\pm i\phi_3}. \end{cases} \quad (28)$$

For the interaction energy we take the form

$$V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{2e^2}{r_3} - \frac{e^2}{r_{13}} - \frac{e^2}{r_{23}}. \quad (29)$$

It is easy to see that $V_{0f} = 0$ owing to the opposite symmetry of ψ_0 and ψ_f . The contribution then comes only from the exchange integrals which give the following expression for the matrix element V'_{0f} ,

$$V'_{0f} = N_1 N_2 e^2 u^{-\frac{1}{2}} (I_1 I_2 + J_1 J_2),$$

where
$$I_1 = \frac{4\pi}{3} \int_0^\infty r^{s+q} \exp \left\{ -\left(\frac{z}{s} + \frac{\zeta'}{q} \right) r \right\} dr, \quad (30)$$

$$J_1 = 4\pi \int_0^\infty r^{s+p} \exp \left\{ -\left(\frac{z}{s} + \frac{\zeta}{q} \right) r \right\} dr, \quad (31)$$

$$I_2 = \int_{r_3 < a} d\mathbf{r}_3 \int_{r_1^{s+p-2} r_3^{p-1} \left(\frac{1}{r_3} - \frac{1}{r_{13}} \right)} \exp \left\{ ikr_3 \cos \theta_3 - \left(\frac{z}{s} + \frac{\zeta}{p} \right) r_1 - \frac{\zeta}{p} r_3 \right\} d\mathbf{r}_1, \quad (32)$$

$$J_2 = \int_{r_3 < a} d\mathbf{r}_3 \int_{r_1^{s+q-2} r_3^{p-1} \left(\frac{1}{r_3} - \frac{1}{r_{13}} \right)} \cos^2 \theta_1 \times \exp \left\{ ikr_3 \cos \theta_3 - \left(\frac{z}{s} + \frac{\zeta'}{q} \right) r_1 - \frac{\zeta}{p} r_3 \right\} d\mathbf{r}_1. \quad (33)$$

To evaluate I_2 , J_2 , we use the expansion (23). In this case

$$\begin{aligned} \gamma_0 &= 0, \quad r_1 < r_3, \\ &= \frac{1}{r_3} - \frac{1}{r_1}, \quad r_1 > r_3, \end{aligned}$$

and γ_n is as before, for values of $n \geq 1$. Proceeding as before and using the fact that

$$\cos^2 \theta = \frac{1}{3}[1 + 2P_2(\cos \theta)],$$

the integrals may be reduced to the forms

$$I_2 = \frac{16\pi^{\frac{3}{2}}}{(2k)^{\frac{1}{2}}} \int_a^\infty dr_3 \int_{r_1}^\infty r_1^{s+p} r_3^{p+\frac{1}{2}} J_{\frac{1}{2}}(kr_3) \exp \left\{ -\left(\frac{z}{s} + \frac{\zeta}{p} \right) r_1 - \frac{\zeta}{p} r_3 \right\} \left(\frac{1}{r_3} - \frac{1}{r_1} \right) dr_1, \quad (34)$$

$$\begin{aligned} J_2 &= \frac{16\pi^{\frac{3}{2}}}{3(2k)^{\frac{1}{2}}} \int_a^\infty dr_3 \int_{r_1}^\infty r_1^{s+q} r_3^{p+\frac{1}{2}} J_{\frac{1}{2}}(kr_3) \exp \left\{ -\left(\frac{z}{s} + \frac{\zeta'}{q} \right) r_1 - \frac{\zeta}{p} r_3 \right\} \left(\frac{1}{r_3} - \frac{1}{r_1} \right) dr_1 \\ &\quad - \frac{32\pi^{\frac{3}{2}}}{(2k)^{\frac{1}{2}}} \int_a^\infty dr_3 \int_0^\infty r_1^{s+q} r_3^{p+\frac{1}{2}} J_{\frac{1}{2}}(kr_3) \exp \left\{ -\left(\frac{z}{s} + \frac{\zeta'}{q} \right) r_1 - \frac{\zeta}{p} r_3 \right\} \gamma_2(r_1 r_3) dr_1. \end{aligned} \quad (35)$$

These integrals may be evaluated numerically and the calculation of $m(a)$ completed.

(3) $n(a)$. Here ψ_0 represents the negative ion and is given by (28). For ψ_f we have

$$\psi_f = \chi(\mathbf{r}_3) \phi_0(\mathbf{r}_1, \mathbf{r}_2), \quad (36)$$

where
$$\phi_0(\mathbf{r}_1, \mathbf{r}_2) = N_3 r_1^{s-1} r_2^{s-1} \exp \left[-\frac{z_0}{s} (r_1 + r_2) \right], \quad (37)$$

and represents the ground state of the mercury atom. We have for V the expression (29). Proceeding as before V_{0f} may be reduced to the form

$$V_{0f} = 32\pi^2 N_2 N_3 e^2 \left(\frac{\pi}{2ku} \right)^{\frac{1}{2}} K \int_a^\infty dr_3 \int_0^\infty r_3^{q+\frac{1}{2}} r_2^{p+s} \gamma_0(r_2, r_3) J_{\frac{3}{2}}(kr_3) \\ \times \exp \left\{ - \left(\frac{z_0}{s} + \frac{\xi}{p} \right) r_2 - \frac{\xi'}{q} r_3 \right\} dr_2, \quad (38)$$

where
$$K = 4\pi \int_0^\infty r^{s+p} \exp \left\{ - \left(\frac{z_0}{s} + \frac{\xi}{p} \right) r \right\} dr. \quad (39)$$

The exchange integral is given by

$$V'_{0f} = \frac{16\pi^{\frac{3}{2}} N_2 N_3 e^2}{3(2ku)^{\frac{1}{2}}} K \int_a^\infty dr_3 \int_0^\infty r_3^{q+\frac{1}{2}} r_2^{p+s} \gamma_1(r_2, r_3) J_{\frac{3}{2}}(kr_3) \\ \times \exp \left\{ - \left(\frac{z_0}{s} + \frac{\xi'}{q} \right) r_2 - \frac{\xi}{p} r_3 \right\} dr_2. \quad (40)$$

These may be evaluated numerically, and $n(a)$ may then be calculated from the expression (4b).

The values of the integrals which we obtain by making the problem spherically symmetrical will be higher, and will decrease less rapidly with a than those we should obtain by using the correct form (16) for $\chi(r)$. This may be seen in the following way. The values of the integrals depend on the extent of the overlap of the regions where the wave functions ψ_0 and ψ_f are appreciable. ψ_f will be small outside a sphere $r_3 = R$, say. In the spherical case ψ_0 will only be appreciable outside a sphere $r_3 = a$. For a plane surface, however, ψ_0 will only be large over part of this region. When we average over all angles for the direction of motion of the "free" electron, the correct form should give values smaller by a factor which will be given approximately by the ratio of the extents of the effective regions in the two cases. We shall estimate this and use it as a correction.

It might be possible to carry out the calculation for an ion at the centre of a hemispherical cavity in a plane surface, for which under certain conditions an exact form for $\chi(r)$ is obtainable (cf. Lennard-Jones and Goodwin 1937). The considerable increase in length of an already complicated calculation, which this would involve, does not seem worth while at the moment, as the values obtained by the simplified model are as good as we can expect when we consider the other approximations we have made.

The values obtained for the transition probabilities $p(a)$, $m(a)$, $n(a)$, are shown in fig. 2. Using these values we may calculate the function $N(V, W)$. This function is shown in fig. 3 for $V = 200$ eV.

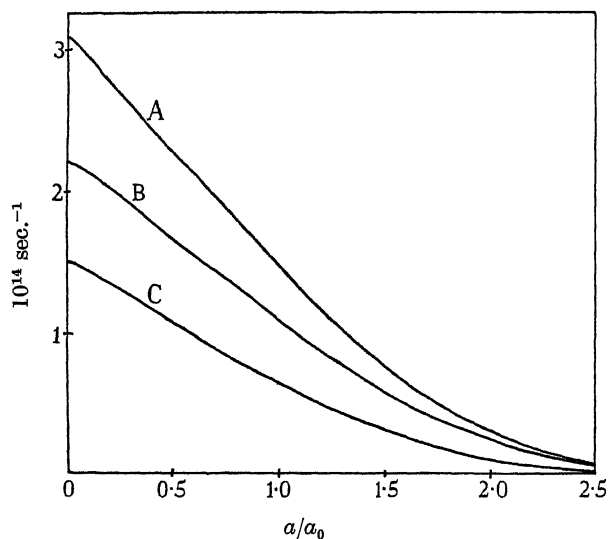


FIG. 2. Transition probabilities $p(a)$, $m(a)$, $n(a)$ for mercury at a nickel surface. A, $n(a)/2$. B, $p(a)/10$. C, $m(a)$.

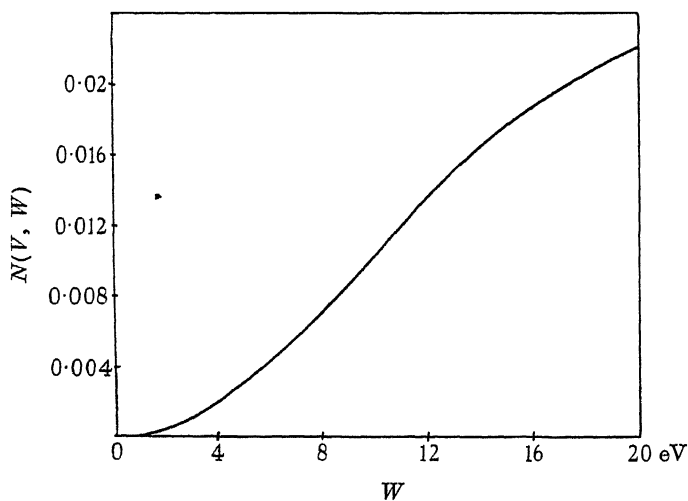


FIG. 3. $N(V, W)$, the probability of a positive ion, incident with energy V and reflected with energy W , forming a negative ion. $V = 200$ eV.

Probability of formation of Hg^- ions at a Ni surface

Before proceeding further we must obtain some information as to the value to take for the mean energy $W_0(V)$ in the distribution function $G(V, W)$. From Arnot and Milligan's experimental curves giving the energy distribution of the negative ions formed from Hg^+ ions we may obtain the function $G(V, W) N(V, W)$. The most accurate measurements are those giving the energy distribution of negative ions retaining a considerable part of the kinetic energy of the positive ions from which they were formed. We shall choose the value of W_0 for $V = 200$ eV so that our calculated curve for $G(V, W) N(V, W)$ fits the experimental curve as closely as possible. In fig. 4

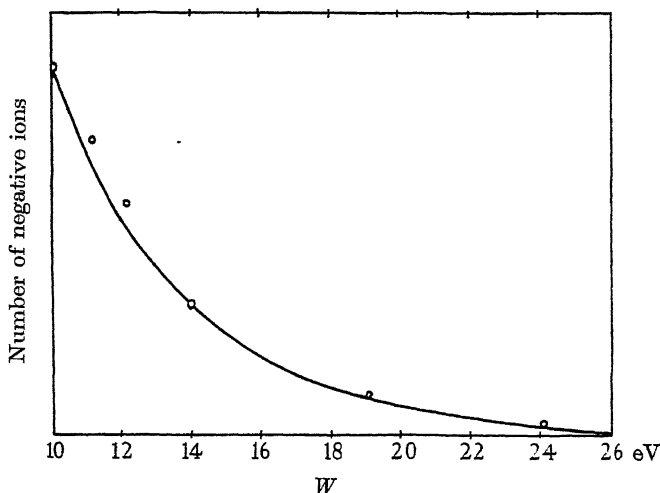


FIG. 4. Energy distribution of negative ions, calculated with $W_0 = 2.75$ eV for $V = 200$ eV. The circles are Arnot and Milligan's experimental points, fitted to the curve at $W = 10$ eV.

the calculated curve for $V = 200$ eV, taking the value 2.75 eV for W_0 , is shown along with Arnot and Milligan's experimental points. It will be seen that a good fit is possible, so we shall use this value of W_0 to calculate the average probability of formation $\bar{N}(V)$ by means of the expression (14). The value we obtain is 1.4×10^{-3} for $V = 200$ eV. This agrees as well as can be expected with the experimental value 0.64×10^{-3} .

Arnot and Milligan have given a curve for the distribution of the high-energy negative ions only for $V = 200$ eV. It is not therefore possible to find W_0 for different values of V by the method used for $V = 200$ eV, and so determine theoretically the variation of $\bar{N}(V)$ with V . Using the value of W_0 for $V = 200$ eV we may, however, estimate W_0 for other values of V by

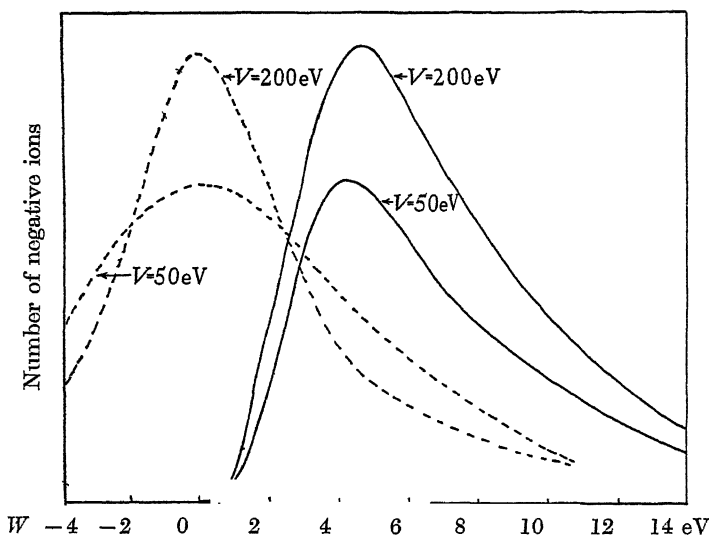
using the experimental curve for the variation of $\bar{N}(V)$ with V . These values of W_0 may then be used to calculate energy distributions of the negative ions for various values of V , which may then be compared with the experimental curves for the ions retaining a small part of their original kinetic energy which, although they have a distinctive shape, are not sufficiently accurate to determine W_0 owing to the low resolving power of the apparatus for slow ions. In Table I we give some values of W_0 obtained in this way and the corresponding values of $\bar{N}(V)$ calculated relative to the theoretical value for $V = 200$ eV. It will be seen that W_0 decreases with V but less rapidly than if it were proportional to V . This we should expect as a result of the increase of the effective cross-section for scattering by surface atoms as V decreases.

TABLE I

Energy of incident positive ions V in eV	Mean energy after reflexion $W_0(V)$ in eV	Calculated probability of formation of Hg^- ions $\bar{N}(V) \times 10^{-3}$
200	2.75	1.4
100	2.3	.86
50	1.65	.36
20	1.2	.09

Using the above values of W_0 we have calculated the energy distribution for negative ions which leave the surface with small kinetic energy, for $V = 50$ and $V = 200$ eV. These are shown in fig. 5 along with the corresponding experimental curves of Arnot and Milligan. When we compare the theoretical and experimental curves, we see that, although they have the same general shape the maxima do not occur at the same values of W . From the experimental curves it would appear that many of the observed negative ions have very little excess energy over that acquired in the electric field which accelerates them away from the surface at which they are formed. This is only so for mercury. The curves for several other gases for which energy distributions of the negative ions have been measured resemble those we have obtained theoretically for mercury, the maxima being shifted out to small positive values of W . The discrepancy between the theoretical and experimental curves for mercury almost certainly arises from our neglect of the effect of the "image" force on the motion of the ions. In the calculation we are concerned with the normal velocity of the ions close to the surface, whereas in the experiments the velocity is observed at a large distance from the surface. Since the range of the "image" force is large compared with the distance from the surface within which transitions are probable, a negative ion moving away from the surface after formation will be retarded and will

lose a certain amount of its kinetic energy. The quantity W used in the calculations refers to the kinetic energy of the ions close to the surface and so will not be the same as the observed kinetic energy of the negative ions. The difference may amount to several electron volts. The effect of this will be that the maxima in the experimental curves will appear displaced to the left relative to the corresponding maxima in the theoretical curves. This will be more marked in mercury than in the lighter gases where the reflected ions retain more of their original kinetic energy.



Kinetic energy of negative ions (excess over accelerating potential in experimental curves)

FIG. 5. Energy distribution of slow negative ions.

— Calculated curves.

--- Arnot and Milligan's curves (fitted to have same maximum height).

It is clear that negative ions having small velocities when formed would almost certainly be neutralized before moving far from the surface, so that those observed with zero excess energy must have been subsequently retarded. The effect of the "image" force seems to offer the most simple explanation of this retardation. When we take into account the low resolving power of the apparatus for slow ions this effect seems adequate to explain the difference between the theoretical and experimental curves for the energy distribution of the slow negative ions. We may therefore say that the main features of the experiments on the formation of Hg^- ions from Hg^+ ions are explained by assuming the negative ions to be formed by process (3).

Formation of negative ions from metastable atoms

By a slight modification of the calculation we have given, we may obtain the chance that a metastable atom formed in the vapour will become a negative ion on colliding with a metal surface. This can be seen to be of the same order of magnitude as for a positive ion of the same energy, since the latter has a fairly large chance of capturing an electron and becoming a metastable atom before reaching the surface. Experiments on the number of Hg^- ions formed with different pressures of mercury vapour indicate that not many of the observed Hg^- ions are formed from metastable atoms. This simply means that most of the metastable atoms striking the metal surface are too slow to be effective.

Formation of negative ions by simultaneous capture of two electrons from a metal

We shall now consider process (4). We shall show that the chance of this process taking place is small compared with (3). Let \mathbf{r}_1 and \mathbf{r}_2 represent the co-ordinates of the two electrons involved in the transition and $\chi_1(\mathbf{r}_1)$, $\chi_2(\mathbf{r}_2)$ the wave functions for their "free" motion in the metal. Let \mathbf{r}_3 represent the co-ordinates of the electron in the outer shell of the mercury positive ion, and let $\phi(\mathbf{r}_3)$ be the wave function for the normal state of the ion. To the same approximation as we used before the wave function for the initial state of the system will be

$$\psi_0 = \chi_1(\mathbf{r}_1) \chi_2(\mathbf{r}_2) \phi(\mathbf{r}_3). \quad (41)$$

If we express the wave function for the negative ion as a product of three "single-particle" wave functions as before it will be of the form

$$\psi_f = \psi_1(\mathbf{r}_1) \psi_2(\mathbf{r}_2) \psi_3(\mathbf{r}_3). \quad (42)$$

Neglecting the interaction of the "free" electrons we have for the interaction energy under which the transition takes place

$$V = e^2 \left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{r_{13}} - \frac{1}{r_{23}} \right). \quad (43)$$

The matrix element for the transition is then

$$V_{0f} = e^2 \iiint \chi_1(\mathbf{r}_1) \chi_2(\mathbf{r}_2) \phi(\mathbf{r}_3) \cdot V \cdot \psi_1(\mathbf{r}_1) \psi_2(\mathbf{r}_2) \psi_3(\mathbf{r}_3) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3, \quad (44)$$

which may be expressed in the form

$$\begin{aligned}
 V_{0f} = & e^2 \int \chi_1(\mathbf{r}_1) \psi_1(\mathbf{r}_1) d\mathbf{r}_1 \int \chi_2(\mathbf{r}_2) \phi(\mathbf{r}_3) \left(\frac{2}{r_2} - \frac{1}{r_{23}} \right) \psi_2(\mathbf{r}_2) \psi_3(\mathbf{r}_3) d\mathbf{r}_2 d\mathbf{r}_3 \\
 & + e^2 \int \chi_2(\mathbf{r}_2) \psi_2(\mathbf{r}_2) d\mathbf{r}_2 \int \chi_1(\mathbf{r}_1) \phi(\mathbf{r}_3) \left(\frac{2}{r_1} - \frac{1}{r_{13}} \right) \psi_1(\mathbf{r}_1) \psi_3(\mathbf{r}_3) d\mathbf{r}_1 d\mathbf{r}_3.
 \end{aligned} \tag{45}$$

This would be zero if the "continuous" wave functions $\chi_1(\mathbf{r}_1)$, $\chi_2(\mathbf{r}_2)$ were accurately orthogonal to the corresponding functions $\psi_1(\mathbf{r}_1)$, $\psi_2(\mathbf{r}_2)$. The transition probability will therefore be small and would require a second order calculation using more accurate wave functions than the simple products we have used.*

3. FORMATION OF HYDROGEN NEGATIVE IONS

The H^- ions which were observed by Arnot were shown to be mainly produced by impact of H_2^+ on a tungsten surface. The number of H^+ ions available was, however, small, and no information was obtainable on the formation of H^- ions from H^+ . The latter process is particularly interesting as condition (c) for the formation of the intermediate excited atom is not satisfied. It is clearly desirable that further experiments on the formation of H^- ions should be made with separated beams of protons and H_2^+ ions. If the H^- ions are formed, as Arnot has suggested, from H^+ ions produced by the dissociation of H_2^+ , the main difficulty is to see how excited hydrogen atoms rise so as to enable process (3) to take place. They cannot be formed by neutralization of the H^+ ions as we have seen that the condition (c) for this is not satisfied. Let us consider various possibilities.

An H_2^+ ion may be neutralized before reaching the surface by capturing an electron from the metal. From the potential energy curves for the H_2 molecule and H_2^+ ion it can be seen that condition (c) is satisfied only for the $1^1\Sigma$, $1^3\Sigma$ states of the molecule. If an electron is captured in to the $1^3\Sigma$ state the molecule will dissociate into two normal hydrogen atoms. It is clear that no excited atoms will be formed in this way. If they are formed from the normal H_2^+ ion or its products, excitation must take place on impact with the surface. This does not appear likely as hydrogen may only be excited with difficulty by impact with heavy atoms even at considerably higher voltages than those with which we are concerned (Massey and Smith 1933). Another possibility is that some of the H_2^+ ions reaching the surface

* I am indebted to Dr H. S. W. Massey for pointing this out, and for other valuable suggestions.

may be in excited states. Such states are known (O. W. Richardson 1934) and their energies are such that an electron may be captured from the metal to form an excited state of the H_2 molecule which on dissociation would give an excited H atom.

The only alternative to process (3) is that of simultaneous capture of two electrons. It is just possible that this process is sufficiently probable in the case of hydrogen to explain the observed probability of formation which is only 0.016 of that for mercury ions of the same energy. Further experiments are necessary to decide between these possibilities. We shall calculate, however, the probability of formation of an H^- ion from an excited H atom, as the knowledge of this will be useful in deciding between the above alternatives.

We shall require the transition probabilities $m(a)$, $n(a)$, for formation and loss of the negative ion. We shall also calculate the transition probability $p(a)$ for the formation of an excited H atom for a surface for which (c) is satisfied. We shall use this to estimate the probability of formation of an H^- ion by impact of a proton on such a surface.

(1) $m(a)$, $n(a)$. The wave functions which we require are simply those for the ground state and the degenerate first excited states of the hydrogen atom, and that for the H^- ion. For the latter we take the form

$$\psi_- = N_- e^{-\mu(r_1+r_2)}, \quad (46)$$

where $\mu = 0.73/a_0$. The calculation may be carried out in exactly the same way as for mercury. The values of $m(a)$, $n(a)$ for $\phi = 3$ eV will not differ very much from those calculated for Ni with $\phi = 5.03$ eV.

(2) $p(a)$. For $\phi = 5.03$ eV, $p(a)$ will be zero as the condition (c) is not satisfied. We shall calculate $p(a)$ for a surface for which $\phi = 3$ eV. The integrals we require have been given by Massey (1930), and $p(a)$ may be evaluated on inserting appropriate values for the various constants.

The values we obtain for $m(a)$, $n(a)$ with $\phi = 5.03$ eV and for $p(a)$ with $\phi = 3$ eV are shown in fig. 6. These are considerably larger than the corresponding values for mercury. The effective quantities in determining the probability of formation are, however, $p(a) \div v$, etc., and these will be a good deal smaller for corresponding energies.

The probability that an excited hydrogen atom, which leaves the surface with energy W , will become a negative ion would be obtained by putting $M_0 = 1$ in (13), if we were to neglect the chance of the excited atom ejecting an electron from the metal on falling into its ground state. With hydrogen however, there will be a considerable chance of this taking place, as the

available energy is 5.1 eV. Let $e(a)$ be the transition probability for this taking place. We may then take it into account by replacing equations (8) by

$$\left. \begin{aligned} v \frac{dM'}{da} &= -m(a) M' - e(a) M', \\ v \frac{dN'}{da} &= -n(a) N' + m(a) M'. \end{aligned} \right\} \quad (47)$$

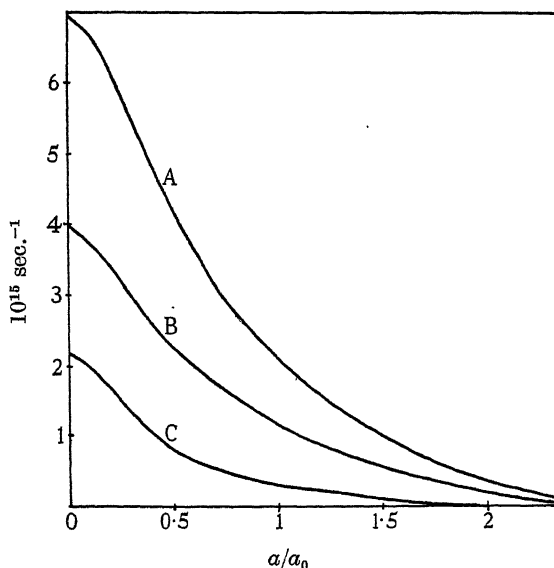


FIG. 6. Transition probabilities $p(a)$, $m(a)$, $n(a)$ for atomic hydrogen.
A, $p(a)$ for $\phi = 3.0$ eV. B, $m(a) \times 5$ for $\phi = 5.03$ eV. C, $n(a)$ for $\phi = 5.03$ eV.

$e(a)$ may be obtained from Massey's calculation (1930) for fairly large values of a , and is found to be about ten times as great as $m(a)$. Using this value, and solving the equations (59) as before we may calculate the probability $N_e(W)$ for the formation of a negative ion from an excited atom which leaves the surface with energy W . This is shown in fig. 7.

It will be seen that for corresponding energies $N_e(W)$ is a good deal larger than for mercury. This is due mainly to the smaller probability of loss of the negative ions formed, as they are moving much faster than the heavy mercury ions.

For hydrogen it is clear that no simple form of the function $G(V, W)$ will give the observed energy distribution of the negative ions, which has two distinct maxima (Arnot 1937). These may be due to formation by two different processes or simply to the fact that for the light hydrogen ions there is pronounced reflexion with certain energies. In any case we should expect

the energy distribution to be more complicated for the light ions as they retain a much larger fraction of their original energy on being reflected from the surface.

In order to obtain the probability of formation of an H^- ion from an H_2 ion by the process we have considered, we must multiply $N_e(W)$ by the chance that excited H atoms will be formed from the incident H_2^+ ions. It can be seen from fig. 7 that this must be about 10^{-3} in order to obtain the observed value 1.04×10^{-5} for the probability of formation of H^- ions from 200 volt H_2^+ ions. This value is probably small enough to be accounted for by assuming that the excited atoms are formed by dissociation of excited H_2^+ ions.

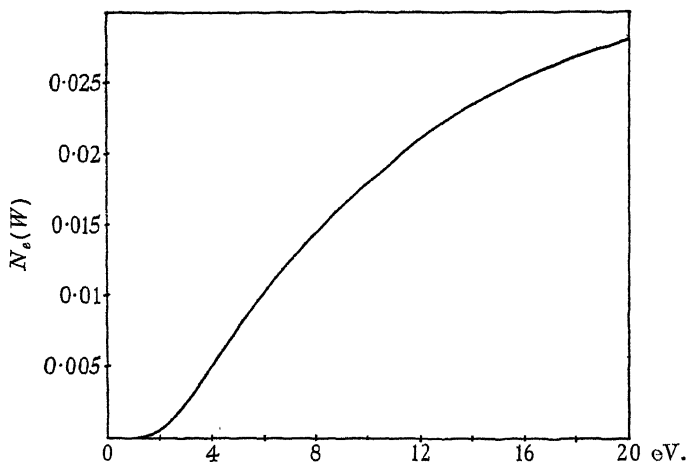


FIG. 7. $N_e(W)$, probability of formation of a negative ion from an excited H atom which leaves a Ni surface with kinetic energy W .

Formation of H^- ions from protons

We should expect the probability of formation of H^- ions from protons to be high for a surface for which $\phi < 3.37$ eV, as then the excited atoms may be formed by capture of electrons from the metal. Using the calculated value of $p(a)$ for $\phi = 3.0$ eV we may estimate the probability that an incident proton will reach the surface as an excited H atom by solving equations similar to (7) but taking into account the effect of $e(a)$. For 140 V protons the value we obtain is 0.2. The probability that such an excited atom will form a negative ion after reflexion will be approximately the same as we have calculated for $\phi = 5.03$ eV. Using this value we may estimate the probability of an incident proton being reflected as a negative ion. This will not vary much with the energy of the reflected ions, provided this is not very small (see fig. 7). The value we obtain for the average probability of formation

of H^- ions by 140 V protons at a surface for which $\phi < 3.37$ eV is about 4×10^{-3} . A surface with a sufficiently low work-function would therefore provide an effective source of hydrogen negative ions but may be impracticable to use.

It would be very interesting if a comparison could be made between surfaces for which ϕ is greater and less than 3.37 eV using a beam of protons. When ϕ is greater than this value we should expect very few H^- ions as they could only be formed by simultaneous capture of two electrons, assuming that hydrogen atoms are not likely to be excited on impact with the surface. An experiment of this type would therefore give information on the probability of such double transitions taking place.

Similar calculations may be carried out for the other polyatomic gases in which Arnot has studied the formation of negative ions. More experiments are needed, as we have seen, in order to clarify the processes taking place and until further information is available for hydrogen it does not seem worth while extending the calculations to include the other gases.

It is a pleasure to acknowledge my indebtedness to Dr F. L. Arnot for keeping me informed on the results of his experiments, and for many helpful discussions.

SUMMARY

Various processes of formation of negative ions at metal surfaces, and the conditions under which they may take place are discussed. Detailed calculations are given for the conversion of Hg^+ ions into Hg^- ions at a nickel surface. The positive ions are assumed to be neutralized by capturing an electron from the metal, and to form excited atoms which subsequently capture another electron on falling into their ground state. This process is shown to account for the experimental results of Arnot and Milligan (1936). The calculated probability of formation by this process is 1.4×10^{-3} for 200 V Hg^+ ions striking a nickel surface normally. The process of simultaneous capture of two electrons from the metal is shown to be improbable.

A negative ion formed near a metal surface for which the work function is greater than the electron affinity of the corresponding atom or molecule will have a considerable chance of being neutralized through one of its electrons passing into an unoccupied level in the metal. It follows that slow positive ions or metastable atoms will be ineffective as a source of negative ions.

The formation of atomic negative ions from molecular positive ions is discussed. Calculations for the formation of H^- ions from H_2^+ ions and from

protons are given. The value obtained for the probability of formation of H^- ions from 140 V protons striking a surface with work function less than 3.37 eV is about 4×10^{-3} . It is shown that H^- ions will not be formed from protons at a surface with work function greater than 3.37 eV unless simultaneous capture of two electrons takes place. For formation of H^- ions from H_2^+ the observed probability is only 1.04×10^{-5} for 200 V H_2^+ ions striking a nickel surface (Arnot 1937). This low value is shown to be due to the small probability of formation of the intermediate excited H atoms by dissociation of the incident ions. Clearly further experiments, using separated beams of protons and H_2^+ ions, are required in order to clarify the processes taking place when atomic negative ions are formed from molecular positive ions.

REFERENCES

- Arnot and Milligan 1936 *Proc. Roy. Soc. A*, **156**, 538.
Arnot, F. L. 1937 *Proc. Roy. Soc. A*, **158**, 137.
Bloch and Bradbury 1935 *Phys. Rev.* **48**, 689.
Compton and Langmuir 1930 *Rev. Mod. Phys.* **2**, 175.
Emeleus and Sayers 1938 *Proc. Roy. Irish Acad.* **44 A**, No. 7.
Fox and Bowie 1933 *Phys. Rev.* **44**, 345.
Glocker 1934 *Phys. Rev.* **46**, 111.
Lennard-Jones and Goodwin 1937 *Proc. Roy. Soc. A*, **163**, 101.
Massey, H. S. W. 1930 *Proc. Camb. Phil. Soc.* **26**, 386.
— 1931 *Proc. Camb. Phil. Soc.* **27**, 460.
Massey and Smith 1933 *Proc. Roy. Soc. A*, **142**, 142.
— — 1936 *Proc. Roy. Soc. A*, **155**, 472.
Oliphant 1929 *Proc. Roy. Soc. A*, **124**, 228.
— 1930 *Proc. Roy. Soc. A*, **127**, 373.
Oppenheimer 1928 *Phys. Rev.* **32**, 361.
Richardson, O. W. 1934 "Molecular Hydrogen and its Spectrum."
Yale Univ. Press.
Slater 1930 *Phys. Rev.* **36**, 57.
Smith, R. A. 1936 *Proc. Camb. Phil. Soc.* **32**, 482.
-

Thermal effects on bodies in an air stream

By W. F. HILTON, PH.D., B.Sc., A.R.C.Sc., D.I.C.

Of the Aerodynamics Department, N.P.L.

(Communicated by Ernest F. Relf, F.R.S.)

1. INTRODUCTION

Very little has been published on the mathematical theory of heat transfer between an unheated solid body and the fluid streaming past it, and the only experimental work seems to be that by H. H. Suplee (1909) giving the temperature drop shown by a thermoelectric wire stretching axially along a diverging nozzle.

E. Pohlhausen (1921) shows theoretically that the temperature (T_1) taken up by a flat plate placed endwise in an air stream of initial temperature (T) should be given by

$$T - T_1 = -\frac{1}{8} \frac{u^2}{c_p} \beta(\sigma), \quad (1)$$

where u is the velocity of the stream outside the boundary layer, c_p the specific heat at constant pressure and $\sigma = \mu c_p / K$ (μ and K being the viscosity and thermal conductivity of air, respectively). The function $\beta(\sigma)$ is tabulated by Pohlhausen (1921) and is given by

$$\beta(\sigma) = 2\sigma \int_0^\infty \exp\left(-\sigma \int_0^\xi \zeta d\xi\right) \left[\int_0^\xi \left(\frac{d^2\zeta}{d\xi^2}\right)^2 \exp\left(\int_0^\xi \zeta d\xi\right) d\xi \right] d\xi,$$

where ζ is the function of ξ tabulated by Blasius (1908), representing the laminar boundary layer on a flat plate.

Equation (1) given by Pohlhausen is not in a suitable form for examination by the present method, and the necessary modifications are made in the next section.

2. APPLICATION OF POHLHAUSEN'S EQUATION TO THE PRESENT WORK

For the purposes of the present paper, three points must be borne in mind when considering Pohlhausen's equation, viz.

- (1) It applies only to a flat plate.
- (2) It is only true in the absence of heat conduction in the fluid or in the plate. This condition is violated by the presence of shock waves.

(3) The temperature (T) of the air stream cannot be measured directly, in any type of wind tunnel, including that used for these experiments (the 1 ft. high-speed wind tunnel at the National Physical Laboratory). In this type of tunnel, an annular jet of compressed air causes a drop in pressure in the tunnel and thus induces atmospheric air to enter a contracting intake, where it is accelerated until it reaches the working section. Since no external work has been done on this air, the energy of its motion past the working section must have been obtained from its own internal heat energy. Thus the air at the working section is cooler than the external air by an amount equal to the adiabatic cooling.

The temperature of the air at the working section of a wind tunnel cannot be measured by inserting a thermometer, since the bulb of the thermometer will stop the air to some extent, thus causing the air to regain as heat energy some of its energy of motion.

It is, however, a simple matter to calculate the cooling due to an adiabatic expansion of the air. If the air temperature be T_0 at the intake, and T at the working section, then

$$T = T_0 - \frac{u^2}{2c_p}. \quad (2)$$

If the expansion were not adiabatic, there would be a loss of total head, as measured by an open-ended pitot tube. On making such a measurement in the empty tunnel against external atmospheric pressure, it is found that such a loss does indeed occur, but is of negligible proportion, being less than 0.1 in. of water at a tunnel speed corresponding to 8 in. of mercury, i.e. less than 0.1 %. We may thus assume that equation (2) gives for all practical purposes the temperature T in the free air stream in the tunnel.

Eliminating T between equations (1) and (2) we have

$$\delta T \equiv T_0 - T_1 = \frac{u^2}{2c_p} \left[1 - \frac{\beta(\sigma)}{4} \right]. \quad (3)$$

Now, δT may be found by inserting one calibrated thermocouple in the surface of the body under test, and the other in the air intake to the tunnel, and measuring the potential difference produced at any given wind speed.

3. TESTS ON FLAT PLATES

Equation (3) gives the temperature difference between the surface of a thin plate with a laminar boundary layer, and the air at the tunnel intake before acceleration past the model. In order to verify this equation experimentally, certain precautions were taken.

First, in order to minimize the transference of heat from one part of the flat plate to another, the models were made of Tufnol, which is at the same time strong, and a good insulator, both electrical and thermal. Difficulty was experienced with testing very thin flat plates, due to deformation, and in more extreme cases, destruction of the model by the aerodynamic forces.

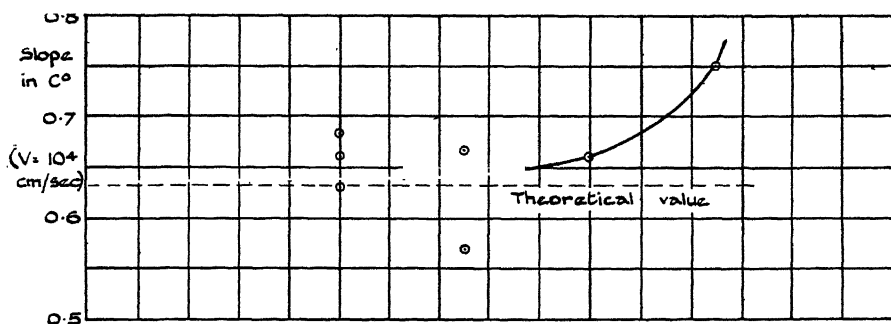


FIG. 1.

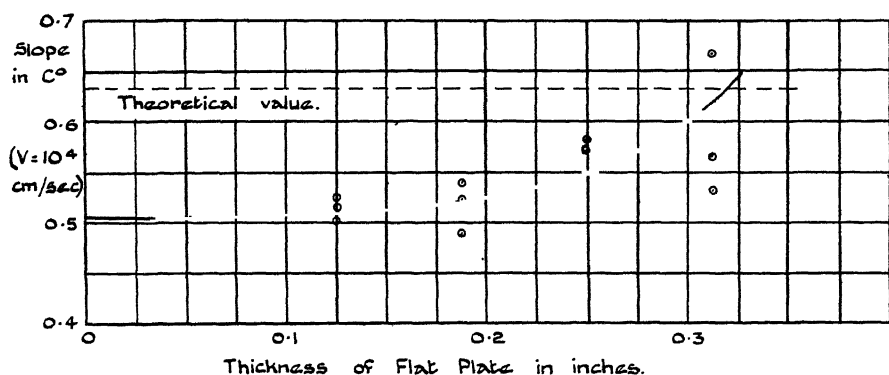


FIG. 2.

Extrapolation of flat plate results to zero thickness.

FIG. 1. Thermocouple nearer leading edge.

FIG. 2. Thermocouple nearer trailing edge.

On this account, a series of four different thicknesses was used, ranging from 1/8 in. to 5/16 in., and the results were extrapolated to zero thickness. The chord in every case was 2 in. and the span the full diameter of the wind tunnel (12 in.). Both leading and trailing edges were semi-elliptical, the major axis being three times the thickness (the minor axis) in each case.

The thermocouple was set flush in the surface, one-third of the chord from the leading edge. By reversing the model the conditions at two-thirds of the chord could be explored. The thermocouple was slightly offset from

the centre of the span, since the thermal conditions at the centre of the tunnel might be unsteady, due to the mixing of air of slightly different temperatures from the six sides of the hexagonal intake.

On plotting temperature difference against V^2 (where V is the tunnel speed), a linear law was found to exist, as is suggested by the theory. It will be noticed from fig. 3, however, that the curves do not always pass

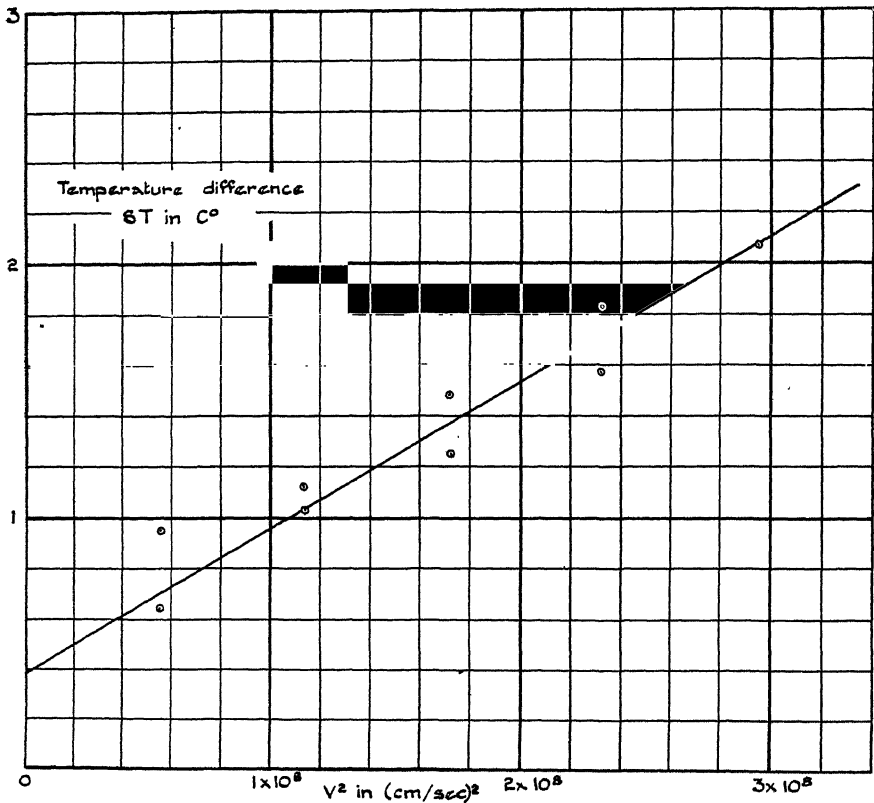


FIG. 3. Typical set of readings for determining slopes plotted in figs. 1 and 2. Flat plate $\frac{3}{16}$ in. thick with thermocouple near leading edge.

through the origin as one would expect, but often show a temperature difference of as much as 1° C. at zero speed. It is thought that this was due to the intake thermocouple being in an air stream of slightly different initial temperature from the thermocouple on the flat plate. The steadiest readings were obtained early in the morning, with all major sources of heat in the building turned off.

The slope of the straight line through each set of readings was measured, and plotted against the thickness of the flat plate, and the

collected results are shown in figs. 1 and 2. The curve through these points has been extrapolated to zero thickness to give the mean slope α (expressed in $^{\circ}\text{C}$ for a speed $V = 10^4$ cm./sec.) for a mathematically thin flat plate.

The mean slopes thus found are $\alpha = 0.636$ with the thermocouple one-third of the chord from the leading edge, and $\alpha = 0.504$ for the thermocouple two-thirds of the chord from the leading edge. Substitution of the following numerical values (supplied by the Physics Department of the National Physical Laboratory) in equation (3) gives a theoretical value for α .

V = air speed over flat plate = u = tunnel speed.

$$c_p = 0.2396 \text{ cal./g./}^{\circ}\text{C.}$$

$$\mu = 18.2 \times 10^{-5} \text{ g./cm./sec.}$$

$$K = 5.73 \times 10^{-5} \text{ cal./g./}^{\circ}\text{C/sec.}$$

$$\sigma = \mu c_p / K = 0.761.$$

$$(1 - \beta(\sigma)/4) = 0.127 \text{ from Pohlhausen's paper, and hence}$$

$$\alpha = \frac{10^8}{2c_p J} \times 0.127 = 0.633.$$

It will be seen from this that the experimental value is in close agreement with the theory near the leading edge, but is lower by some 26 % near the trailing edge. This is probably due to the flow over the front half of the flat plate being laminar, where Pohlhausen's theory is applicable and to a transition to turbulent flow occurring between the alternative thermocouple positions. Consideration of the local Reynolds number at these positions, which ranges from 0.8×10^5 to 1.8×10^5 near the leading edge and double this near the trailing edge, lends weight to this explanation, as does the lack of turbulence in the wind tunnel stream. In addition we have the fact that α is less for the turbulent flow, corresponding to an increased heating effect (since cooling is considered positive throughout this paper), which is to be expected from the extra dissipation of energy in turbulent flow.

Thus it would seem that equations (1) and (3) for a flat plate in laminar flow are substantiated within the limits of experimental accuracy.

It will be seen on reference to figs. 1 and 2 that the effect of thickness is small up to $\frac{1}{4}$ in. thick, and that in fig. 1 the experimental points lie fairly evenly on both sides of the theoretical value for zero thickness. The mean experimental value ($\alpha = 0.636$) is only 0.003 above the theoretical value ($\alpha = 0.633$), whilst the average experimental error would seem to be much larger than this, say 0.05.

The difference between the curves for leading and trailing edges is very marked, however, and much too great to be ascribed to experimental error. The theoretical value may itself be slightly in error, since it is difficult to measure accurately the thermal conductivity of air, and an error of 1 % in K gives an error of 5 % in the theoretical value of the temperature.

4. TESTS ON A STREAMLINE BAR AT NON-COMPRESSIBILITY SPEEDS

Whilst equation (3) is only strictly true for thin flat plates, it is of interest to examine the possibility of its application to laminar boundary layers on other shaped bodies, and in particular to the front part of an aerofoil. We may assume that the boundary layer will be but little affected by gentle curvature in the direction of the wind, and we now consider u to be the velocity just outside the boundary layer, and not the tunnel velocity (V). (The two were equal in the case of a flat plate.) Hence by making the substitution $u = nV$ in equation (3), where the factor (n) expresses the increase in wind speed due to the aerofoil, we have

$$\delta T = \frac{n^2 V^2}{2C_p} \left[1 - \frac{\beta(\sigma)}{4} \right]. \quad (4)$$

It should be remarked that n varies continuously round the aerofoil section under test. Equation (4) is in a form directly suitable for experimental checking in the high speed tunnel. A thermocouple was soldered flush with the surface of the streamline section under test, as shown in fig. 5, and the difference of e.m.f. between this thermocouple, and another placed in the slow moving air at the intake, was measured.

On plotting temperature difference against the square of the wind speed (as is done in figs. 4–7) it was found that a linear law again holds, up to compressibility speeds.

This is in qualitative agreement with equation (4), and experimental values of n have been found from the slopes of these straight lines, but the results are not given here, since the model was of steel, a good conductor of heat, and hence the values found for n were more in the nature of averages for the whole section, rather than values appropriate to the point at which the thermocouple was situated. It is hoped to make a model from heat insulating material at a later date, and to continue these experiments under more appropriate conditions.

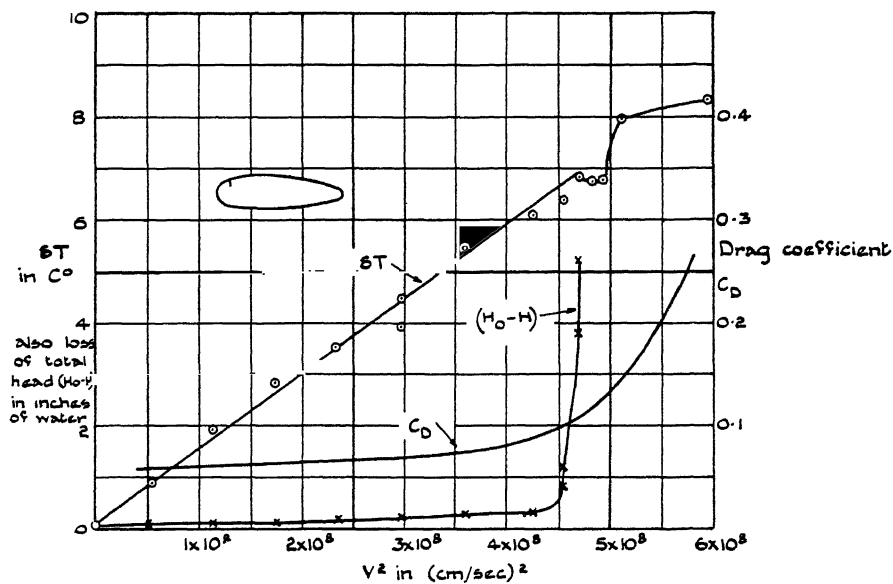


FIG. 4. Streamline bar at 0° incidence.

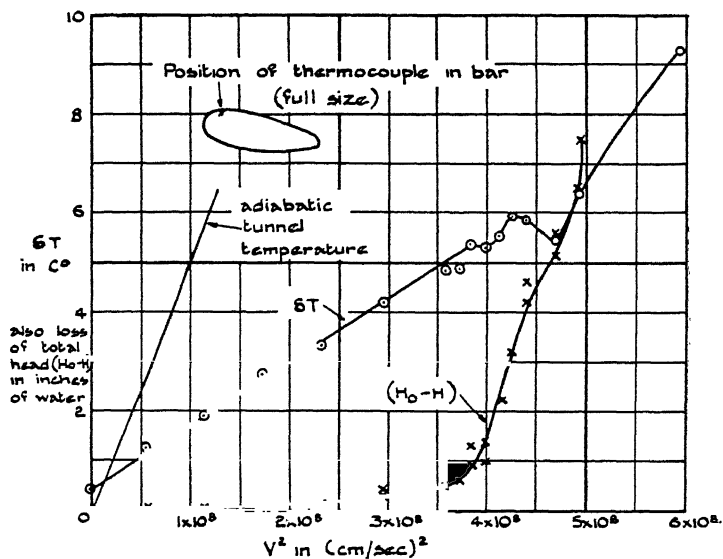


FIG. 5. Streamline bar at 5.5° incidence.

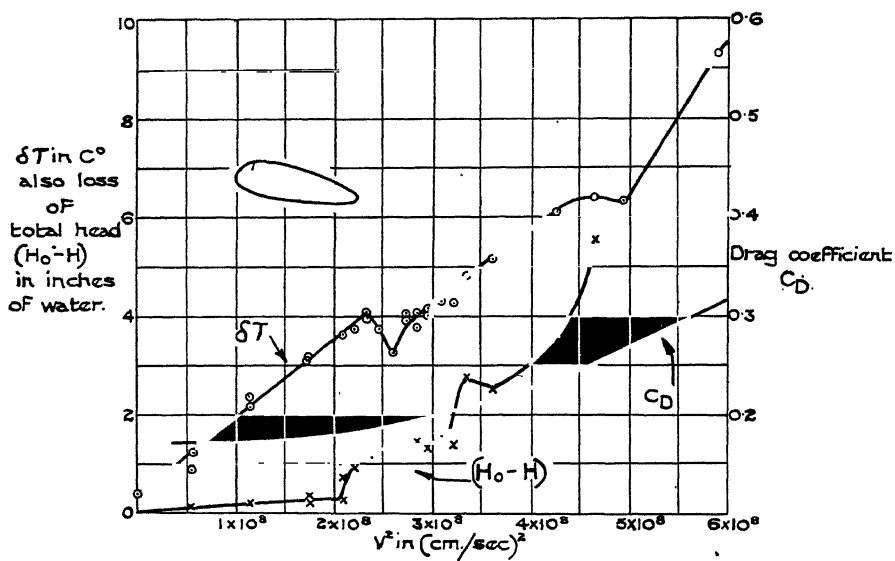


FIG. 6. Streamline bar at 8° incidence.

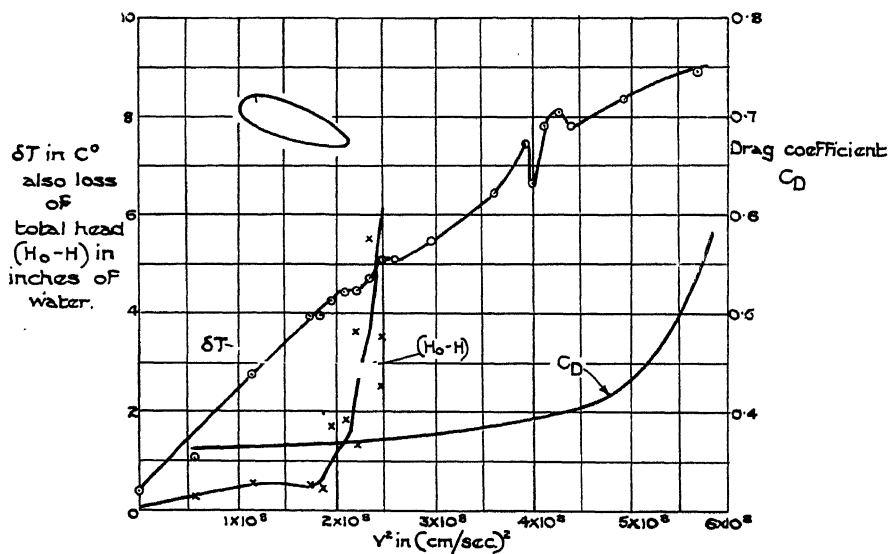


FIG. 7. Streamline bar at 15° incidence.

5. TESTS ON A STREAMLINE BAR AT COMPRESSIBILITY SPEEDS

The linear relation existing between the temperature difference and the square of the speed breaks down at high speeds, where a small speed range is usually found in which the temperature difference δT decreases with increases of speed. On examination of the readings as plotted, it might be thought that a smooth curve could be drawn through the points, ignoring the small part with negative slope. This is not in accordance with experiment, since the effect can be observed by leaving the potentiometer switched on and raising or lowering the wind speed continuously. At a critical speed, the galvanometer spot is seen to stop moving regularly across the screen, retrogress for some distance, and then resume its original motion.

Thus the small region of negative slope appears to have a real existence, although it might easily be overlooked if the readings were taken at fairly wide intervals. At even higher speeds a further linear law is found to hold good, the curve having roughly the same slope as at low speeds, but this part of the curve is often broken by further inflexions.

In the derivation of Pohlhausen's equation for a flat plate, the effects of heat conduction in the fluid are assumed to be negligible, so that the inflexions found in the curves as V/a approaches unity may be taken to indicate that the effect of compressibility is no longer small, and that a shock wave has been formed, with its associated conduction of heat. The presence of a shock wave may be detected experimentally, a very sensitive method being to use a pitot tube to measure the total head of the air stream. At low speeds in the empty tunnel, the pitot reads atmospheric pressure (see § 2). At low speeds in the wake of an aerofoil, there is the well-known loss of total head, corresponding to the "drag" forces on the aerofoil. This friction wake is very narrow, being usually less than the thickness of the aerofoil. Accordingly, if a pitot tube is placed just outside the friction wake and the speed increased, no appreciable reading will be obtained until a shock wave is formed. Since there is a loss of total head in passing through a shock wave, there will be a sudden rise in the curve of loss of total head against wind speed at the precise speed at which a shock wave is first formed. This constitutes a very delicate test for the existence of shock waves, and in figs. 4-7 are shown a series of such curves, plotted for convenience of reference on the same speed scale as the thermocouple readings.

It will be seen from these curves that an inflexion in the temperature curve coincides with the inception of a shock wave. It should be remembered that throughout this paper positive δT represents cooling, so that a portion of curve with negative slope represents a sudden heating, corre-

sponding to the heating in the shock wave due to the loss of kinetic energy. Actually, a body is always warmer than the air flowing steadily past it, even

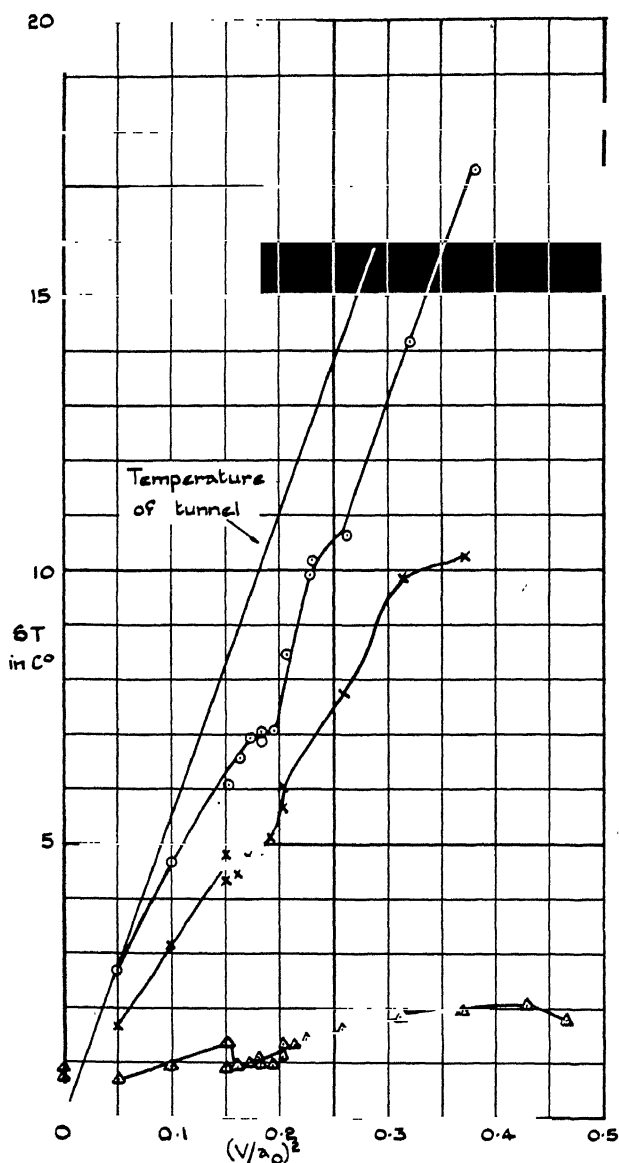


FIG. 8. Temperature variation of circular cylinder with speed.

Position of thermocouple: \triangle at front ($a_0 = 34460$ cm./sec.); \times at side ($a_0 = 34390$ cm./sec.); \circ at back ($a_0 = 34180$ cm./sec.).

at low speeds, but this warming is always masked in the high speed tunnel by the previous adiabatic cooling of the air. For a body moving through still air this warming would increase linearly with the square of the speed and at the "shock stall" the temperature would rise even more rapidly. The initial adiabatic cooling of the air is represented on the diagrams by a straight line joining the origin and the point ($V^2 = 10^8$, $\delta T = 4.97^\circ \text{C}$). The heating of the body is found from the vertical difference between this line and the curve measured in the tunnel.

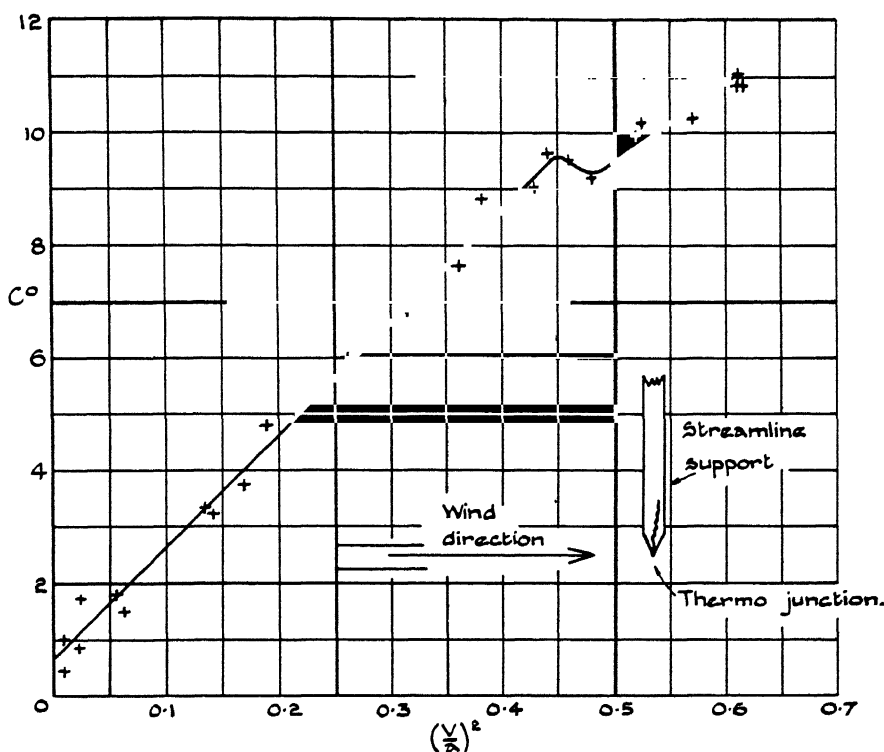


FIG. 9. General Electrical Company thermocouple, extending halfway across tunnel.

The existence of shock waves can also be detected by the fairly sudden increase in the drag of the aerofoil at the shock stall. Some drag readings were taken on the section actually used, and are shown in figs. 4–7. These will be seen to be in agreement with the results obtained by other methods, but the measurement of drag is not such an accurate method of determining the shock stall speed as either of those already discussed.

Thus it would seem that the present thermocouple method is a quick and convenient test for the presence of a local shock-wave at any particular

spot on a solid boundary, and the method may prove valuable when aircraft speeds are limited by the formation of shock waves, since it provides a method for determining which part of the machine is giving rise to the increased drag associated with shock waves.

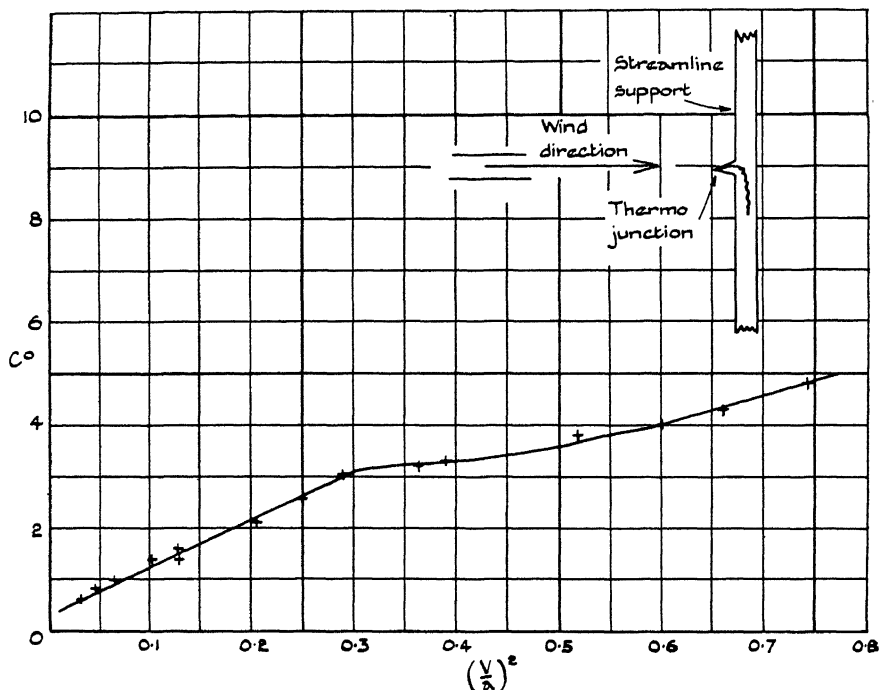


FIG. 10. General Electrical Company thermocouple, extending right across tunnel.

6. THE CIRCULAR CYLINDER AND OTHER SHAPES

A circular cylinder made of Tufnol placed across stream was also tested. From the results plotted in fig. 8 it will be seen that the front of a cylinder registers very nearly the adiabatic compression temperature of the air stream. Had it done so exactly, the temperature difference would have been zero. The thermocouple placed at the back, however, registers within 10 % the true temperature of the moving air stream in which it is placed. Thus on a circular cylinder are to be found almost the extreme ranges of temperature possible on a body in a moving air stream. The circular cylinder is therefore a most unsatisfactory shape for the thermometry of moving fluids, since the temperature taken up by the element will be a function of the conductivity of the thermometric element, and of the temperature dis-

tribution round the cylinder. If the cylindrical shape is unavoidable, the axis should be placed along the wind. Trouble is also to be expected from the critical speeds of cylinders. In the case of the 1/4 in. diameter cylinder employed in these experiments, the critical speed comes in the middle of the speed range, at about $V/a = 0.4$.

In figs. 9 and 10 some results are given for two forms of thermocouple which were tested at the request of the General Electric Company. These results also emphasize the importance of the form of a thermometric element to be used at high speeds.

The best thermometric form would seem to be a short flat plate, for which the calibration is known in advance, since the flow would in general be laminar.

SUMMARY

Pohlhausen's equation for the temperature at the surface of a flat plate with a laminar boundary layer has been verified to within the limits of experimental accuracy. The theoretical temperature slope ($\alpha = 0.633$) is in excellent agreement with the experimental value found near to the leading edge ($\alpha = 0.636$) where the flow is probably laminar. The result near to the trailing edge ($\alpha = 0.504$) is consistent with the flow being turbulent in this region. The linear relation between δT and V^2 breaks down at high speeds, due to the formation of shock waves, and an inflexion occurs in the curve. By measurements of total head in the wake, and by force measurements on the balance, it has been verified that this inflexion is caused by a shock wave. Thus thermocouples may be used to detect the presence of shock waves, and it is suggested that by placing thermocouples on models of aeroplanes, or even on actual aeroplanes in flight, the parts giving rise to shock waves could easily be found, and modified to obtain a higher top speed.

Tests were also made on a steel streamline bar at various incidences and Pohlhausen's equation has been extended to apply to this case, but a rigid verification could not be effected, due to the unknown effect of heat conduction along the metal streamline bar. It is hoped to carry out further tests on aerofoil sections made of heat insulating material, and to compare the values of n found by the present method with those found by direct experiment. Agreement is to be expected over the nose of the aerofoil where the flow will be laminar, but a difference should be found behind the point of transition to turbulent flow.

A circular cylinder placed across stream was found to be a most unsatisfactory shape for a thermometric element, since its temperature varies from nearly the adiabatic compression temperature at the front, to nearly the free stream temperature at the rear. The ideal method of measuring the temperature of a moving fluid would be to move a thermometric element of any shape with the same velocity as the fluid, but since this would usually be impracticable, the best method is to use a stationary flat plate, with the sensitive area in the region of laminar boundary layer flow.

REFERENCES

- Blasius 1908 *Z. Math. Phys.* 56, 1.
Pohlhausen 1921 *Z. angew. Math. Mech.* 1, 115–121.
Supplee 1909 “The Gas Turbine”, Chap. v.
-

Statistical theory of superlattices with long-range interaction

I. General theory

BY J. S. WANG

(Communicated by R. H. Fowler, F.R.S.—Received 30 May 1938)

1. INTRODUCTION

For alloys in which the numbers of two kinds of atoms A and B are equal the statistical theory of superlattices was first developed by Bethe (1935), assuming interaction only between atoms which are nearest neighbours to one another. The theory was then extended by Peierls (1936) to the case of unequal concentrations of components still retaining the assumption of interaction between nearest neighbours, and by Chang (1937) to include the interaction between next nearest pairs of atoms in the case of equal concentrations of components. It is desirable to have a more general theory which includes long-range interactions of a general type and is applicable both to equal and unequal concentrations of components, but the interaction energy must fall off sufficiently rapidly with increasing distance to make the contribution of very distant atoms negligible so that the shape and

the dimensions of the crystal block do not influence the state of the superlattice. For example, if the interaction energy is of the form $V(r) = Ar^{-n}$, then n should be greater than 3 in order that the integral

$$\int_0^\infty V(r) r^2 dr$$

may be finite.

A general theory of this kind has already been developed for a similar problem in adsorption (Wang 1938). The problem of adsorption is simpler than that of superlattices for two reasons. First, the former is a two-dimensional problem while the latter is a three-dimensional one. Second, there is only one kind of atom in the adsorption case and therefore every lattice site is equivalent to every other, whereas in the superlattice case there are two kinds of atoms and hence two different kinds of lattice sites, usually called α sites and β sites respectively. This second difference between the problem of adsorption and that of superlattices is the one which makes the direct application of the adsorption theory impossible, because the device of replacing the contribution of distant adsorbed atoms by that of a continuous uniform distribution no longer works for the superlattice, since it would lead to a sort of amalgamation of the two kinds of lattice sites. This difficulty is overcome in the present theory as is explained in § 2.

The general theory developed here is applicable to any interaction energy of the form $V(r)$ subject to the condition that $\int_0^\infty V(r) r^2 dr$ is finite, and to any crystal lattice in which the total number of sites at any given distance from a given α site is the same as that from a β site. The important case of an interaction energy following the law of inverse power of distance is worked out for the three lattices of the cubic system—the simple cubic, body-centred and face-centred—all of which satisfy the above requirement for the lattice structure. It should be mentioned that the theory is restricted to the case where the number of A atoms is equal to that of α sites and the number of B atoms to that of β sites. The effect of a slight variation in the composition of the alloy consisting of two kinds of atoms A and B is not investigated here.†

2. THE CONFIGURATIONAL PARTITION FUNCTION

The problem of superlattices is solved completely if the partition function of the alloy crystal is known. To construct the partition function we shall confine ourselves to that part which is contributed purely by the geo-

† Such a problem has been studied by C. E. Easthope (1937).

metrical arrangements of the atoms and is called the configurational partition function. It is assumed that the configurational energy is independent of the lattice vibrations.

Let a_1 be the distance between nearest neighbours in the lattice, a_2 that between the next nearest, etc. Denote by V_ν^{AA} , V_ν^{BB} , V_ν^{AB} the interaction energies between pairs AA , BB , AB at a distance a_ν apart, and by M_ν^{AA} , M_ν^{BB} , M_ν^{AB} the numbers of such pairs. The configurational energy is then

$$E = \sum_\nu \{M_\nu^{AA}V_\nu^{AA} + M_\nu^{BB}V_\nu^{BB} + M_\nu^{AB}V_\nu^{AB}\}. \quad (1)$$

According to the general theory of statistical mechanics the configurational partition function is

$$P = \sum_M \gamma_M e^{-E/kT},$$

where γ_M is the number of arrangements of the atoms for a given set of values M_ν^{AA} , M_ν^{BB} , M_ν^{AB} ($\nu = 1, 2, \dots$), and the summation is taken over all such possible sets.

Suppose that the total number of lattice sites is N , the number of α sites is rN and that of β sites is $(1-r)N$. We assume that the number of A atoms is rN , that of B atoms is $(1-r)N$, and the total number of atoms is N . We assume also that the number of sites situated at a distance a_ν from an α site is the same as that from a β site and denote it by z_ν . Let Z_ν be the sum of z_ν for all the N sites of the crystal,† that contributed by α sites being rZ_ν , and that by β sites $(1-r)Z_\nu$. Now Z_ν/N multiplied by the number rN of A atoms is equal to the total number of pairs of atoms at a distance a_ν apart involving A atoms, each AA pair among them being counted twice. Therefore

$$rZ_\nu = 2M_\nu^{AA} + M_\nu^{AB}.$$

Similarly
$$(1-r)Z_\nu = 2M_\nu^{BB} + M_\nu^{AB}.$$

Substituting in (1) we obtain

$$E = \sum_\nu M_\nu^{AA}V_\nu + \text{const.}, \quad (2)$$

where
$$V_\nu = V_\nu^{AA} + V_\nu^{BB} - 2V_\nu^{AB}. \quad (3)$$

The additive constant in (2) is independent of the atomic arrangements and will be put equal to zero, since the zero of energy can be chosen arbitrarily.

† For a site near the boundary the number of sites situated at a distance a_ν from it is less than z_ν and therefore $Z_\nu < Nz_\nu$. But when N is large the percentage error involved in taking $Z_\nu = Nz_\nu$ is small.

With the expression (2) for the energy the partition function can be constructed by specifying the numbers of pairs AA only, viz. the numbers M_{ν}^{AA} . We can therefore ignore the presence of B atoms altogether, and concentrate our attention solely on the A atoms. Henceforth V_{ν} will, for the sake of simplicity, be called the interaction energy between two A atoms at a distance a_{ν} apart, and it is important to bear this convention in mind.

We evaluate the partition function approximately by treating in detail the distribution of A atoms among a central site and its z nearest neighbours (which form the first shell), and replacing the influence of A atoms in the outer sites on the first shell by the introduction of a suitable parameter for each A atom in the first shell and the influence on the central site by that of an average distribution of A atoms among the outer sites compatible with the particular state of the superlattice in question. Since we now do not fix the number of A atoms in the group of $z+1$ sites we must, following the usual procedure of statistical mechanics, introduce a selector variable† ξ for each A atom and determine it by the condition that the average probability of occupation of the central site by an A atom calculated by using the approximate partition function is equal to the actual average corresponding to a total number rN of A atoms. The approximate partition function thus constructed depends, in general, on whether the central site is an α or a β site. Accordingly we distinguish two cases:

(i) *Central α site*

Let n_0 be the number of A atoms in the central site, so that n_0 is either 1 or 0 according as the central site is occupied by an A atom or not. Let $n_1 = n'_1 + n''_1$ be the number of A atoms in the z first shell sites, n'_1 being the number of A atoms in the α sites of the first shell and n''_1 that in the β sites. Let m_{ν} be the number of AA pairs among the $n_0 + n_1$ A atoms at a distance a_{ν} apart. Denote the energy of interaction of an A atom situated at the central α site with an average distribution of A atoms outside the first shell by U_{α} . The exact meaning of the average distribution will be made clear in §6 when we come to the actual evaluation of U_{α} . If we introduce for the influence of the outside A atoms on the first shell a factor ϵ_{α} for each of the n'_1 A atoms in the α sites and a factor ϵ_{β} for each of the n''_1 A atoms in the β sites (the selector variable ξ for A atoms in the first shell is absorbed in the parameters ϵ_{α} and ϵ_{β}), we obtain the following expression for the approximate partition function

$$Q_{\alpha} = \sum \gamma (\xi e^{-U_{\alpha}/kT})^{n_0} \epsilon_{\alpha}^{n'_1} \epsilon_{\beta}^{n''_1} \prod_{\nu} \eta_{\nu}^{m_{\nu}},$$

† Peierls (1936) calls this an *a priori* probability.

where γ is the number of arrangements for a given set of values n_0, n'_1, n''_1, m_ν , and η_ν is given by

$$\eta_\nu = e^{-V_\nu/kT} \quad (\nu = 1, 2, \dots). \quad (4)$$

The quantity m_1 can be split into two terms

$$m_1 = n_0 n_1 + m'_1,$$

where m'_1 is the number of AA pairs among the n_1 A atoms in the first shell at a distance a_1 apart (i.e. nearest neighbours). Then Q_α becomes

$$Q_\alpha = F(\epsilon_\alpha, \epsilon_\beta) + \xi e^{-U_\alpha/kT} F(\eta_1 \epsilon_\alpha, \eta_1 \epsilon_\beta), \quad (5)$$

where
$$F(x, y) = \sum \gamma x^{n_1'} y^{n_1''} \eta_1^{m_1'} \prod_{\nu \geq 2} \eta_\nu^{m_\nu}. \quad (6)$$

The average values of n_0, n'_1, n''_1 are given by the usual formulae

$$\bar{n}_0 = \xi \frac{\partial}{\partial \xi} \log Q_\alpha, \quad \bar{n}'_1 = \epsilon_\alpha \frac{\partial}{\partial \epsilon_\alpha} \log Q_\alpha, \quad \bar{n}''_1 = \epsilon_\beta \frac{\partial}{\partial \epsilon_\beta} \log Q_\alpha.$$

Introducing the notations

$$\left. \begin{aligned} F &= F(\epsilon_\alpha, \epsilon_\beta), & F_\eta &= F(\eta_1 \epsilon_\alpha, \eta_1 \epsilon_\beta), \\ \phi &= \phi(\epsilon_\alpha, \epsilon_\beta) = \epsilon_\alpha \frac{\partial}{\partial \epsilon_\alpha} \log F, & \phi_\eta &= \phi(\eta_1 \epsilon_\alpha, \eta_1 \epsilon_\beta), \\ \psi &= \psi(\epsilon_\alpha, \epsilon_\beta) = \epsilon_\beta \frac{\partial}{\partial \epsilon_\beta} \log F, & \psi_\eta &= \psi(\eta_1 \epsilon_\alpha, \eta_1 \epsilon_\beta), \end{aligned} \right\} \quad (7)$$

we obtain

$$\left. \begin{aligned} \bar{n}_0/(1 - \bar{n}_0) &= \xi e^{-U_\alpha/kT} F_\eta/F, \\ \bar{n}'_1 &= \bar{n}_0 \phi_\eta + (1 - \bar{n}_0) \phi, & \bar{n}''_1 &= \bar{n}_0 \psi_\eta + (1 - \bar{n}_0) \psi. \end{aligned} \right\} \quad (8)$$

(ii) Central β site

Let all the n 's and m 's denote the same things as before. Denote the energy of interaction of an A atom situated at the central β site with an average distribution of A atoms outside the first shell by U_β . If we introduce for the influence of the outside A atoms on the first shell a factor ζ_α for each of the n'_1 A atoms in the α sites of the first shell and a factor ζ_β for each of the n''_1 A atoms in the β sites, we obtain the following expression for the approximate partition function

$$\begin{aligned} Q_\beta &= \sum \gamma (\xi e^{-U_\beta/kT})^{n_0} \zeta_\alpha^{n_1'} \zeta_\beta^{n_1''} \prod_\nu \eta_\nu^{m_\nu} \\ &= G(\zeta_\alpha, \zeta_\beta) + \xi e^{-U_\beta/kT} G(\eta_1 \zeta_\alpha, \eta_1 \zeta_\beta), \end{aligned} \quad (9)$$

where
$$G(x, y) = \sum \gamma x^{n_1'} y^{n_1''} \eta_1^{m_1'} \prod_{\nu \geq 2} \eta_\nu^{m_\nu}. \quad (10)$$

The expression (10) for G is of exactly the same form as the expression (6) for F . Actually they are different, owing to the different possible sets of values for the n 's and m 's. As all the subsequent formulae involve F and G explicitly, no confusion is possible.

Introducing the notations

$$\theta = \zeta_\alpha \frac{\partial}{\partial \zeta_\alpha} \log G, \quad \chi = \zeta_\beta \frac{\partial}{\partial \zeta_\beta} \log G, \quad (11)$$

and also G_η , θ_η , χ_η analogous to F_η , ϕ_η , ψ_η , we obtain

$$\left. \begin{aligned} \bar{n}_0/(1-\bar{n}_0) &= \xi e^{-U_\beta/kT} G_\eta/G, \\ \bar{n}'_1 &= \bar{n}_0 \theta_\eta + (1-\bar{n}_0) \theta, \quad \bar{n}''_1 = \bar{n}_0 \chi_\eta + (1-\bar{n}_0) \chi. \end{aligned} \right\} \quad (12)$$

The determination of ξ , ϵ_α , ϵ_β , ζ_α , ζ_β in terms of the superlattice order is given in the next section.

3. THE SUPERLATTICE ORDER

Let r_α be the average probability of an α site being occupied by an A atom and $w_\alpha = 1 - r_\alpha$ that of its being occupied by a B atom. Let r_β be the average probability of a β site being occupied by a B atom and $w_\beta = 1 - r_\beta$ that of its being occupied by an A atom. The total number of A atoms is then

$$\{rr_\alpha + (1-r)w_\beta\}N.$$

Equating this to rN we obtain

$$rr_\alpha + (1-r)w_\beta = r,$$

or

$$rw_\alpha = (1-r)w_\beta. \quad (13)$$

This condition also gives the right number for the B atoms, for we have

$$rw_\alpha + (1-r)r_\beta = (1-r)w_\beta + (1-r)r_\beta = 1-r.$$

The degree of superlattice order s is defined by the equation

$$s = (r_\alpha - r)/(1-r). \quad (14)$$

This definition makes $s=1$ for perfect order and $s=0$ for perfect disorder. In the state of perfect order every α site is occupied by an A atom and every β site by a B atom. In the state of perfect disorder every site has, *on the average*, a probability of occupation r for A atoms and $1-r$ for B atoms.

From (13) and (14) we obtain

$$\left. \begin{aligned} r_\alpha &= r + (1-r)s, & w_\alpha &= (1-r)(1-s), \\ r_\beta &= 1-r+rs, & w_\beta &= r(1-s). \end{aligned} \right\} \quad (15)$$

In applying the partition function of the foregoing section to calculate the various probabilities r_α , etc., we again divide the investigation into two parts:

(i) *Central α site*

Here we want to determine ξ , ϵ_α , ϵ_β in terms of s . For this purpose we impose the following conditions

$$\bar{n}_0 = r_\alpha, \quad \bar{n}'_1 = z_{\alpha\alpha}r_\alpha, \quad \bar{n}''_1 = z_{\alpha\beta}w_\beta, \quad (16)$$

where $z_{\alpha\alpha}$ is the number of α sites and $z_{\alpha\beta}$ that of β sites in the first shell. The physical meaning of (16) is quite simple. The first equation expresses that the average probability of occupation of the central α site is equal to the average for all the α sites and it determines the selector variable ξ in terms of s through r_α . The second and third equations express that the average probabilities of occupation of a first shell α site and a first shell β site are the same as the respective values determined by s , and are the equations for fixing the parameters ϵ_α and ϵ_β in terms of s .

Substituting (16) in (8) we obtain

$$r_\alpha/w_\alpha = \xi e^{-U_\alpha/kT} F_\eta / F \quad (17)$$

and
$$z_{\alpha\alpha}r_\alpha = r_\alpha\phi_\eta + w_\alpha\phi, \quad z_{\alpha\beta}w_\beta = r_\alpha\psi_\eta + w_\alpha\psi. \quad (18)$$

(ii) *Central β site*

Here we determine ξ , ζ_α , ζ_β in terms of s by imposing the following conditions

$$\bar{n}_0 = w_\beta, \quad \bar{n}'_1 = z_{\beta\alpha}r_\alpha, \quad \bar{n}''_1 = z_{\beta\beta}w_\beta, \quad (19)$$

where $z_{\beta\alpha}$ is the number of α sites and $z_{\beta\beta}$ that of β sites in the first shell. Substituting (19) in (12) we obtain

$$w_\beta/r_\beta = \xi e^{-U_\beta/kT} G_\eta / G \quad (20)$$

and
$$z_{\beta\alpha}r_\alpha = w_\beta\theta_\eta + r_\beta\theta, \quad z_{\beta\beta}w_\beta = w_\beta\chi_\eta + r_\beta\chi. \quad (21)$$

The equation which determines the degree of superlattice order s as a function of the temperature is obtained by eliminating ξ between (17) and (20), namely,

$$\frac{r_\alpha r_\beta}{w_\alpha w_\beta} = \frac{F_\eta G}{F G_\eta} e^{-(U_\alpha - U_\beta)/kT}. \quad (22)$$

Eqns. (15), (18), (21), (22) are the fundamental equations in the present theory. They are of the type of equations which are usually solved by the method of successive approximations. The actual numerical calculations will be given in a subsequent paper.

It should be observed that the quantities $z_{\alpha\alpha}$, etc., in (18) and (21) are connected by two relations and can therefore be expressed in terms of two parameters. The first relation comes from the assumption of equal z_ν for α and β sites in the case of $\nu = 1$, i.e.

$$z_{\alpha\alpha} + z_{\alpha\beta} = z_{\beta\alpha} + z_{\beta\beta} = z_1 = z. \quad (23)$$

The second relation is one between $z_{\alpha\beta}$ and $z_{\beta\alpha}$ obtained by equating the numbers of neighbouring pairs of $\alpha\beta$ sites calculated in two different ways, viz. $rNz_{\alpha\beta} = (1-r)Nz_{\beta\alpha}$. This gives

$$rz_{\alpha\beta} = (1-r)z_{\beta\alpha} = r(1-r)z', \text{ say.} \quad (24)$$

From (23) and (24) we obtain

$$\left. \begin{aligned} z_{\alpha\alpha} &= z - (1-r)z', & z_{\alpha\beta} &= (1-r)z', \\ z_{\beta\alpha} &= rz', & z_{\beta\beta} &= z - rz'. \end{aligned} \right\} \quad (25)$$

4. THE CONFIGURATIONAL ENERGY

Applying the theory of the foregoing sections we find that the average energy of an A atom situated at an α site is

$$U_\alpha + \overline{(n_1)}_{n_0=1} V_1 = U_\alpha + (\phi_\eta + \psi_\eta) V_1$$

and that of an A atom situated at a β site is

$$U_\beta + \overline{(n_1)}_{n_0=1} V_1 = U_\beta + (\theta_\eta + \chi_\eta) V_1.$$

The total configurational energy of the whole alloy is obtained as usual by taking half the sum of the above expressions for all A atoms, the factor one-half being introduced so that each A atom shall be counted once only. Hence

$$E = \frac{1}{2} r N r_\alpha \{U_\alpha + (\phi_\eta + \psi_\eta) V_1\} + \frac{1}{2} (1-r) N w_\beta \{U_\beta + (\theta_\eta + \chi_\eta) V_1\},$$

or

$$E/NV_1 = \frac{1}{2} r r_\alpha (\phi_\eta + \psi_\eta) + \frac{1}{2} (1-r) w_\beta (\theta_\eta + \chi_\eta) + \frac{1}{2} \{r r_\alpha U_\alpha + (1-r) w_\beta U_\beta\} / V_1 \quad (26)$$

For the perfectly ordered state $s = 1$ we have, from (15),

$$r_\alpha = r_\beta = 1, \quad w_\alpha = w_\beta = 0,$$

and from (18) and (21),

$$\phi_\eta = z_{\alpha\alpha}, \quad \psi_\eta = 0; \quad \theta = z_{\beta\alpha}, \quad \chi = 0.$$

Equation (26) then becomes

$$(E)_{s=1}/NV_1 = \frac{1}{2} r \{z_{\alpha\alpha} + (U_\alpha)_{s=1}/V_1\}. \quad (27)$$

5. APPROXIMATION FOR $s = 0$

For temperature at which the superlattice order does not exist we have perfect disorder and $s = 0$. It follows from (15) that

$$r_\alpha = w_\beta = r, \quad r_\beta = w_\alpha = 1 - r. \quad (28)$$

In this case there is no longer a distinction between α and β sites. We must now have, introducing starred symbols,

$$\left. \begin{aligned} \epsilon_\alpha = \epsilon_\beta = \zeta_\alpha = \zeta_\beta = \epsilon^*, \text{ say,} \\ \phi + \psi = \theta + \chi = \phi^*, \quad \phi_\eta + \psi_\eta = \theta_\eta + \chi_\eta = \phi_\eta^* \end{aligned} \right\} \quad (29)$$

and

$$U_\alpha = U_\beta = U^*. \quad (30)$$

Eqn. (22) becomes an identity and (18) and (21) become

$$zr = r\phi_\eta^* + (1 - r)\phi^*. \quad (31)$$

The energy expression (26) becomes

$$E/NV_1 = \frac{1}{2}r(\phi_\eta^* + U^*/V_1). \quad (32)$$

When $T \rightarrow \infty$ all the η 's approach unity and

$$F_\eta \rightarrow F \rightarrow (1 + \epsilon_\infty^*)^z, \quad \phi_\eta^* \rightarrow \phi \rightarrow z\epsilon_\infty^*/(1 + \epsilon_\infty^*),$$

where ϵ_∞^* is the value of ϵ^* at $T = \infty$. But from (31) we have $\phi^* \rightarrow zr$, so that

$$\epsilon_\infty^* = r/(1 - r). \quad (33)$$

The configurational energy approaches the limit

$$E_\infty = \frac{1}{2}r(zr + U^*/V_1)NV_1. \quad (34)$$

6. CALCULATION OF U_α AND U_β

Let $U_{\alpha\alpha}$ be the interaction energy of an A atom situated at an α site with all other α sites, each of them being supposed to be occupied by an A atom, and let $U_{\alpha\beta}$ be the interaction energy of an A atom situated at an α site with all the β sites, each of them being supposed to be occupied by an A atom. By definition U_α is the interaction energy of an A atom situated at a central α site with an average distribution of A atoms outside the first shell. The meaning of the average distribution is that each α site has r_α A atoms and each β site w_β A atoms. It follows that

$$U_\alpha = r_\alpha(U_{\alpha\alpha} - z_{\alpha\alpha}V_1) + w_\beta(U_{\alpha\beta} - z_{\alpha\beta}V_1). \quad (35)$$

Similarly
$$U_\beta = r_\alpha(U_{\beta\alpha} - z_{\beta\alpha}V_1) + w_\beta(U_{\beta\beta} - z_{\beta\beta}V_1), \quad (36)$$

where $U_{\beta\alpha}$ is the interaction energy of an A atom at a β site with all α sites, each of which being supposed to be occupied by an A atom, and $U_{\beta\beta}$ is the interaction energy of an A atom at a β site with all other β sites, each of which being supposed to be occupied by an A atom.

The four quantities $U_{\alpha\alpha}$, etc., like $z_{\alpha\alpha}$, etc., are connected by two relations. The first is

$$U_{\alpha\alpha} + U_{\alpha\beta} = U_{\beta\alpha} + U_{\beta\beta} = U, \text{ say}, \quad (37)$$

which follows immediately from the assumption of equal z , for α and β sites. The second relation is obtained by equating the interaction energies of α sites with β sites calculated in two different ways, viz. $rNU_{\alpha\beta} = (1-r)NU_{\beta\alpha}$. This gives

$$rU_{\alpha\beta} = (1-r)U_{\beta\alpha} = r(1-r)U', \text{ say}. \quad (38)$$

From (37) and (38) we obtain

$$\left. \begin{aligned} U_{\alpha\alpha} &= U - (1-r)U', & U_{\alpha\beta} &= (1-r)U', \\ U_{\beta\alpha} &= rU', & U_{\beta\beta} &= U - rU'. \end{aligned} \right\} \quad (39)$$

These equations are similar to (25) for the z 's.

In terms of the new symbols U and U' (35) and (36) become

$$\left. \begin{aligned} U_\alpha &= r_\alpha(U - zV_1) - (1-r)(r_\alpha - w_\beta)(U' - z'V_1), \\ U_\beta &= w_\beta(U - zV_1) + r(r_\alpha - w_\beta)(U' - z'V_1). \end{aligned} \right\} \quad (40)$$

From (40) we obtain

$$U_\alpha - U_\beta = s\{U - U' - (z - z')V_1\} \quad (41)$$

and

$$rU_\alpha + (1-r)U_\beta = r(U - zV_1) = U^*. \quad (42)$$

It can easily be verified by substituting (28) in (40) that U^* in (42) is the same quantity as given in (30) and is independent of the superlattice order.

The explicit expressions for U_α and U_β when V_ν in (3) is of the form

$$V_\nu = V_1(a_1/a_\nu)^n, \quad (n > 3), \dagger$$

in the case of the three lattices of the cubic system are given below:

† n here must not be confused with the n 's used earlier for numbers of atoms.

(a) *The simple cubic lattice*

Here we have, by definition,

$$\left. \begin{aligned} U_{\alpha\alpha} &= U_{\beta\beta} = V_1 \sum_{\substack{l_1+l_2+l_3 \\ =\text{even}}} (l_1^2 + l_2^2 + l_3^2)^{-\frac{1}{2}n} = A_n'' V_1, \\ U_{\alpha\beta} &= U_{\beta\alpha} = V_1 \sum_{\substack{l_1+l_2+l_3 \\ =\text{odd}}} (l_1^2 + l_2^2 + l_3^2)^{-\frac{1}{2}n} = A_n' V_1, \\ U &= U_{\alpha\alpha} + U_{\alpha\beta} = (A_n' + A_n'') V_1 = A_n V_1, \quad U' = 2A_n' V_1, \end{aligned} \right\} \quad (43)$$

where l_1, l_2, l_3 are positive or negative integers or zero and A_n' and A_n'' are the respective sums for $l_1 + l_2 + l_3$ odd or even. Substituting in (40) and noticing that $r = \frac{1}{2}$, $z = 6$, $z' = 12$, we obtain

$$\begin{aligned} U_{\alpha}/V_1 &= r_{\alpha}(A_n - 6) - (r_{\alpha} - w_{\beta})(A_n' - 6), \\ U_{\beta}/V_1 &= w_{\beta}(A_n - 6) + (r_{\alpha} - w_{\beta})(A_n' - 6). \end{aligned}$$

By (15) with $r = \frac{1}{2}$ these equations reduce to

$$\left. \begin{aligned} U_{\alpha}/V_1 &= \frac{1}{2}(1+s)(A_n - 6) - s(A_n' - 6), \\ U_{\beta}/V_1 &= \frac{1}{2}(1-s)(A_n - 6) + s(A_n' - 6). \end{aligned} \right\} \quad (44)$$

Then (41) and (42) become

$$U_{\alpha} - U_{\beta} = s(6 + A_n'' - A_n') V_1 \quad (45)$$

and

$$U^* = \frac{1}{2}(A_n - 6) V_1. \quad (46)$$

(b) *The body-centred cubic lattice*

Here we have, by definition,

$$\left. \begin{aligned} U_{\alpha\alpha} &= U_{\beta\beta} = V_1 \sum_{\text{all } l} (\frac{1}{2}\sqrt{3})^n (l_1^2 + l_2^2 + l_3^2)^{-\frac{1}{2}n} = (\frac{1}{2}\sqrt{3})^n A_n V_1, \\ U_{\alpha\beta} &= U_{\beta\alpha} = V_1 \sum_{\text{all } l} (\frac{1}{2}\sqrt{3})^n [(l_1 + \frac{1}{2})^2 + (l_2 + \frac{1}{2})^2 + (l_3 + \frac{1}{2})^2]^{-\frac{1}{2}n} = (\frac{1}{2}\sqrt{3})^n B_n' V_1, \\ U &= (\frac{1}{2}\sqrt{3})^n (A_n + B_n') V_1 = B_n V_1, \quad U' = 2(\frac{1}{2}\sqrt{3})^n B_n' V_1. \end{aligned} \right\} \quad (47)$$

Substituting in (40) to (42) and noticing that $r = \frac{1}{2}$, $z = 8$, $z' = 16$, we obtain

$$\left. \begin{aligned} U_{\alpha}/V_1 &= \frac{1}{2}(1+s)(B_n - 8) - s[(\frac{1}{2}\sqrt{3})^n B_n' - 8], \\ U_{\beta}/V_1 &= \frac{1}{2}(1-s)(B_n - 8) + s[(\frac{1}{2}\sqrt{3})^n B_n' - 8], \end{aligned} \right\} \quad (48)$$

$$U_{\alpha} - U_{\beta} = s[8 + (\frac{1}{2}\sqrt{3})^n (A_n - B_n')] V_1, \quad (49)$$

$$U^* = \frac{1}{2}(B_n - 8) V_1. \quad (50)$$

(c) The face-centred cubic lattice

Here we have, by definition,

$$\left. \begin{aligned} U_{\alpha\alpha} &= V_1 \left(\sum_{\text{all } l} + \sum_{\text{all } l} \right) \left[\frac{1}{2} \{ (l_2 + l_3)^2 + (l_3 + l_1)^2 + (l_1 + l_2)^2 \} \right]^{-\frac{1}{2}n} \\ &= V_1 \sum_{\text{all } m} [2(m_1^2 + m_2^2 + m_3^2)]^{-\frac{1}{2}n} = 2^{-\frac{1}{2}n} A_n V_1, \\ U_{\alpha\beta} &= C_n V_1 - U_{\alpha\alpha} = (C_n - 2^{-\frac{1}{2}n} A_n) V_1, \end{aligned} \right\} \quad (51)$$

$$\text{where} \quad C_n = \sum_{\text{all } l} \left[\frac{1}{2} \{ (l_2 + l_3)^2 + (l_3 + l_1)^2 + (l_1 + l_2)^2 \} \right]^{-\frac{1}{2}n}. \quad (51')$$

Then it follows from (37) and (38) that

$$U = C_n V_1, \quad U' = \frac{4}{3} (C_n - 2^{-\frac{1}{2}n} A_n) V_1. \quad (51'')$$

Now $r = \frac{1}{4}$, $z = 12$, $z' = 16$, so that (40) to (42) become

$$\left. \begin{aligned} U_{\alpha}/V_1 &= \frac{1}{4} (1 + 3s) (C_n - 12) - s (C_n - 2^{-\frac{1}{2}n} A_n - 12), \\ U_{\beta}/V_1 &= \frac{1}{4} (1 - s) (C_n - 12) + \frac{1}{3} s (C_n - 2^{-\frac{1}{2}n} A_n - 12), \end{aligned} \right\} \quad (52)$$

$$U_{\alpha} - U_{\beta} = \frac{1}{3} s (12 + 4 \times 2^{-\frac{1}{2}n} A_n - C_n) V_1, \quad (53)$$

$$U^* = \frac{1}{4} (C_n - 12) V_1. \quad (54)$$

The quantities A_n , A'_n , A''_n , B_n , C_n have been evaluated by Lennard-Jones and Ingham (1925), and their numerical values will be used in a subsequent paper.

In conclusion I wish to thank Dr J. K. Roberts for his kind help.

SUMMARY

Bethe's theory of order-disorder in superlattices is extended to cover the case of long-range interactions. The theory is applicable to any interaction energy of the form $V(r)$ subject to the condition $\int_0^{\infty} V(r) r^2 dr = \text{finite}$, and to a general class of lattice structures which include the three lattices of the cubic system.

REFERENCES

- Bethe, H. A. 1935 *Proc. Roy. Soc. A*, **150**, 552-75.
 Chang, T. S. 1937 *Proc. Roy. Soc. A*, **161**, 546-63.
 Easthope, C. E. 1937 *Proc. Camb. Phil. Soc.* **33**, 502-17.
 Lennard-Jones, J. E. and Ingham, A. E. 1925 *Proc. Roy. Soc. A*, **107**, 636-53.
 Peierls, R. 1936 *Proc. Roy. Soc. A*, **154**, 207-22.
 Wang, J. S. 1938 *Proc. Camb. Phil. Soc.* **34**, 238-52.

Statistical theory of superlattices with long-range interaction

II. The simple cubic lattice and the body-centred cubic lattice

BY J. S. WANG

(Communicated by R. H. Fowler, F.R.S.—Received 30 May 1938)

1. GENERAL FORMULAE

The purpose of this paper is to apply the general theory developed in the preceding paper (quoted as I) to the simple cubic lattice and the body-centred cubic lattice. For these two lattices the nearest neighbours of an α site are all β sites and those of a β site are all α sites. The number of A atoms is now equal to that of B atoms and the number of α sites is equal to that of β sites. The structure of these two lattices is such that†

$$z_{\alpha\alpha} = z_{\beta\beta} = 0, \quad z_{\alpha\beta} = z_{\beta\alpha} = z, \quad z' = 2z, \quad (1)$$

where $z = 6$ for the simple cubic lattice and $z = 8$ for the body-centred cubic lattice (cf. I, eqn. (25)). Since $r = \frac{1}{2}$, eqn. (15) in I becomes

$$r_\alpha = r_\beta = \frac{1}{2}(1+s), \quad w_\alpha = w_\beta = \frac{1}{2}(1-s). \quad (2)$$

Now when the central site is an α site all the first shell sites are β sites, and when the central site is a β site all the first shell sites are α sites. Hence we have only one parameter ϵ and one parameter ζ , namely ϵ_β and ζ_α . To simplify the notation we shall write ϵ and ζ instead of ϵ_β and ζ_α and define

$$\left. \begin{aligned} \psi &= \psi(\epsilon) = \epsilon \frac{d}{d\epsilon} \log F(\epsilon), & \psi_\eta &= \psi(\eta_1 \epsilon), \\ \theta &= \theta(\zeta) = \zeta \frac{d}{d\zeta} \log G(\zeta), & \theta_\eta &= \theta(\eta_1 \zeta). \end{aligned} \right\} \quad (3)$$

Then eqns. (18), (21), (22) in I become

$$\frac{1+s}{1-s} = \frac{z-\psi}{\psi_\eta} = \frac{\theta_\eta}{z-\theta}, \quad \left(\frac{1+s}{1-s} \right)^2 = \frac{F_\eta G}{F G_\eta} \eta_1^{2qs}, \quad (4)$$

where

$$2qs = (U_\alpha - U_\beta)/V_1. \quad (5)$$

† The meanings of the various symbols are the same as in I.

When the interaction energy follows an inverse power law we have, from (45) in I for the simple cubic lattice,

$$q = \frac{1}{2}(6 + A_n'' - A_n'), \quad (6a)$$

and from (49) in I for the body-centred cubic lattice

$$q = \frac{1}{2}[8 + (\frac{1}{2}\sqrt{3})^n (A_n - B_n')]. \quad (6b)$$

The quantity q depends only on the law of force and the structure of the lattice and is independent of the degree of the superlattice order.

Since for the two lattices in question the α sites and β sites are on equal footings, the possible sets of values of the n 's and m 's (cf. I, § 2) are the same for a central α site as for a central β site and the functions F and G become identical. Hence

$$G(\zeta) \equiv F(\zeta), \quad \theta(\zeta) \equiv \psi(\zeta). \quad (7)$$

It can be verified from the explicit forms of the function F given in the next section that it satisfies the following relation

$$F(x) = (\eta_1^{-\mu-\frac{1}{2}} x)^z F(\eta_1^{2\mu+1} x^{-1}), \quad (8)$$

where, in the case of an inverse power law for the interaction energy,

$$\mu = -\frac{1}{2}[1 + 4 \times 2^{-\frac{1}{2}n} + 2^{-n}] \quad (9a)$$

for the simple cubic lattice and

$$\mu = -\frac{1}{2}[1 + 3 \times (\frac{4}{3})^{-\frac{1}{2}n} + 3 \times (\frac{8}{3})^{-\frac{1}{2}n} + 2^{-n}] \quad (9b)$$

for the body-centred cubic lattice. From (8) we deduce by differentiation

$$\psi(x) = z - \psi(\eta_1^{2\mu+1} x^{-1}). \quad (10)$$

Using (7) and (10) we obtain

$$\frac{\theta_\eta}{z-\theta} \equiv \frac{\theta(\eta_1 \zeta)}{z-\theta(\zeta)} = \frac{\psi(\eta_1 \zeta)}{z-\psi(\zeta)} = \frac{z-\psi(\eta_1^{2\mu} \zeta^{-1})}{\psi(\eta_1^{2\mu+1} \zeta^{-1})},$$

so that the first set of equations in (4) becomes

$$\frac{1+s}{1-s} = \frac{z-\psi(\epsilon)}{\psi(\eta_1 \epsilon)} = \frac{z-\psi(\eta_1^{2\mu} \zeta^{-1})}{\psi(\eta_1^{2\mu+1} \zeta^{-1})}.$$

A possible solution of this equation is

$$\epsilon \zeta = \eta_1^{2\mu}. \quad (11)$$

It can be shown by direct differentiation that $d\psi/d\epsilon$ is always positive, and therefore $\{z - \psi(\epsilon)\}/\psi(\eta_1\epsilon)$ is a decreasing function of ϵ . It then follows that (11) is the only solution. Substituting (11) in (4) and using (8), we obtain

$$\left. \begin{aligned} \frac{1+s}{1-s} &= \frac{z-\psi}{\psi_\eta}, \\ \frac{1+s}{1-s} &= \eta_1^{-\frac{1}{2}z+qs} \frac{F_\eta}{F}, \end{aligned} \right\} \quad (12)$$

and

the second equation being obtained after the extraction of the square root. These two equations determine ϵ and s as functions of the temperature and constitute the fundamental equations in the present theory.

2. EXPLICIT FORMULAE FOR THE FUNCTION F

In this section the explicit forms of the function F for the two lattices in question are given.

(a) *The simple cubic lattice*

Here with a central α site we have $n'_1 = 0$, $m'_1 = 0$, $m_3 = 0$, and $m_\nu = 0$ for $\nu > 4$. The values of the weight function γ for various values of n''_1 , m_2 , m_4 are given below:

n''_1	0	1	2	2	3	3	4	4	5	6
m_2	0	0	1	0	3	2	5	4	8	12
m_4	0	0	0	1	0	1	1	2	2	3
γ	1	6	12	3	8	12	12	3	6	1

Then the function $F(x)$ becomes

$$\begin{aligned} F(x) = & 1 + 6x + (12\eta_2 + 3\eta_4)x^2 + (8\eta_2^3 + 12\eta_2^2\eta_4)x^3 \\ & + (12\eta_2 + 3\eta_4)\eta_2^4\eta_4x^4 + 6\eta_2^3\eta_4^2x^5 + \eta_2^{12}\eta_4^3x^6. \end{aligned} \quad (13)$$

When the interaction energy follows an inverse power law, since

$$a_2 = \sqrt{2}a_1, \quad a_4 = 2a_1,$$

we have the following expressions for η_2 and η_4 :

$$\eta_2 = \eta_1^{2-\frac{1}{2}\mu}, \quad \eta_4 = \eta_1^{2-\mu}. \quad (14)$$

With the expression (13) for F the relation (8) can be readily verified. In general μ is determined by the equation

$$\eta_1^{-2\mu-1} = \eta_2^4\eta_4, \quad (15)$$

but it reduces to (9a) when the expressions of η_2 and η_4 given in (14) are used.

(b) *The body-centred cubic lattice*

Here with a central α site, we have $n'_1 = 0$, $m'_1 = 0$, $m_4 = 0$, and $m_\nu = 0$, for $\nu > 5$. The values of the weight function γ for various values of n''_1 , m_2 , m_3 , m_5 are given below:

n''_1	0	1	2	2	2	3	3	3	4	4	4	4	4	4	5	5	5*	6	6	6	7	8
m_2	0	0	1	0	0	2	1	0	4	3	3	2	2	0	5	4	3	7	6	6	9	12
m_3	0	0	0	1	0	1	1	3	2	3	2	3	2	6	4	4	6	6	7	6	9	12
m_5	0	0	0	0	1	0	1	0	0	0	1	1	2	0	1	2	1	2	2	3	3	4
γ	1	8	12	12	4	24	24	8	6	8	24	24	6	2	24	24	8	12	12	4	8	1

Then the function $F(x)$ becomes

$$\begin{aligned}
 F(x) = & 1 + 8x + (12\eta_2 + 12\eta_3 + 4\eta_5)x^2 + (24\eta_2^3\eta_3 + 24\eta_2\eta_3\eta_5 + 8\eta_3^3)x^3 \\
 & + (6\eta_2^4 + 8\eta_2^3\eta_3 + 24\eta_2^3\eta_5 + 24\eta_2^2\eta_3\eta_5 + 6\eta_2^2\eta_5^2 + 2\eta_3^4)\eta_3^2x^4 \\
 & + (24\eta_2^2\eta_3 + 24\eta_2\eta_3\eta_5 + 8\eta_3^3)\eta_2^3\eta_3^3\eta_5x^5 + (12\eta_2 + 12\eta_3 + 4\eta_5)\eta_2^6\eta_3^6\eta_5^2x^6 \\
 & + 8\eta_2^9\eta_3^9\eta_5^3x^7 + \eta_2^{12}\eta_3^{12}\eta_5^4x^8.
 \end{aligned} \tag{16}$$

When the interaction energy follows an inverse power law, since

$$a_2 = \frac{2}{\sqrt{3}}a_1, \quad a_3 = \frac{2\sqrt{2}}{\sqrt{3}}a_1, \quad a_5 = 2a_1,$$

we have the following expressions for η_2 , η_3 , and η_5 :

$$\eta_2 = \eta_1^{(\frac{2}{3})^{-in}}, \quad \eta_3 = \eta_1^{(\frac{2}{3})^{-in}}, \quad \eta_5 = \eta_1^{2^{-in}}. \tag{17}$$

One can verify the relation (8) with the expression (16) for F in the same way as before and finds that μ is determined by the equation

$$\eta_1^{-2\mu-1} = \eta_2^3\eta_3^3\eta_5, \tag{18}$$

which yields (9b) when the expressions for η_2 , η_3 and η_5 given in (17) are used.

3. THE CRITICAL TEMPERATURE AND THE SUPERLATTICE ORDER

If we put $\epsilon = \eta_1^\mu$ in (12) and apply (8) and (10) we find

$$\frac{1+s}{1-s} = 1 \quad \text{and} \quad \frac{1+s}{1-s} = \eta_1^{qs},$$

which gives $s = 0$ independent of the temperature. Hence $s = 0$ is a solution of (12) at all temperatures.

When ϵ increases from 0 to η_1^μ , s determined from the first equation of (12) decreases steadily from 1 to 0. This follows from the fact that $\psi(\epsilon)$ is an increasing function of ϵ . To study the value of s determined from the second equation of (12) we shall use the following equation instead

$$\eta_1^{-as} = \eta_1^{-iz} \frac{F_\eta}{F} \frac{\psi_\eta}{z - \psi}, \quad (19)$$

which is obtained by eliminating the factor $(1+s)/(1-s)$ from the two equations of (12). When ϵ increases from 0 to η_1^μ , s determined from (19) increases at first from $-\infty$ to a maximum value (positive) and then decreases steadily to zero.

When ϵ increases from η_1^μ to ∞ , s determined from the first equation of (12) decreases steadily from 0 to -1 , and s determined from (19) decreases at first from 0 to a minimum value (negative) and then increases steadily to infinity. It is easily proved by using (8) and (10) that if s corresponds to a value of ϵ lying between 0 and η_1^μ , then $-s$ corresponds to a value of ϵ lying between η_1^μ and ∞ . The state $-s$ is evidently obtained from the state s by interchanging A and B atoms and therefore does not give anything new. We can therefore confine ourselves to the positive values of s and to the values of ϵ lying between 0 and η_1^μ .

Let $\epsilon = \eta_1^\mu x$. Then the first equation of (12) and eq. (19) are reduced by means of (8) and (10) to

$$\frac{1+s}{1-s} = \frac{\psi(\eta_1^{\mu+1} x^{-1})}{\psi(\eta_1^{\mu+1} x)} \quad (20)$$

and

$$\eta_1^{-as} = x^{-z} \frac{F(\eta_1^{\mu+1} x) \psi(\eta_1^{\mu+1} x)}{F(\eta_1^{\mu+1} x^{-1}) \psi(\eta_1^{\mu+1} x^{-1})}. \quad (21)$$

For a given value of x the above two equations give, in general, two different values of s , which will be denoted by $s_1(x)$ and $s_2(x)$ respectively, s_1 being given by (20) and s_2 by (21). The general shapes of the curves s_1 and s_2 are shown in fig. 1. The curves are drawn only for $0 \leq x \leq 1$, as the values of x in this region correspond to the values of ϵ in the region $0 \leq \epsilon \leq \eta_1^\mu$.

When the temperature is below a certain critical temperature T_c , the two curves s_1 and s_2 intersect at two points, one at $x = 1$, $s = 0$ and the other at a point where s is different from zero. When the temperature is above the critical temperature the two curves have only the single intersection point at $x = 1$, $s = 0$. It is obvious that the condition for the existence of a second

point of intersection is that the absolute value of the slope of s_2 at $x = 1$ should be greater than that of s_1 , that is, since both slopes are negative,

$$\left(\frac{ds_1}{dx}\right)_{x=1} > \left(\frac{ds_2}{dx}\right)_{x=1} \quad (\text{condition for } T < T_c). \quad (22)$$

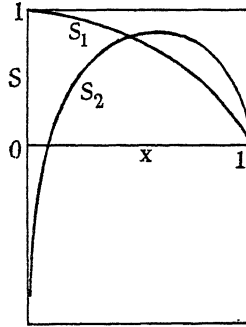


FIG. 1. Curves of $s_1(x)$ and $s_2(x)$, arbitrary scale.

The critical temperature T_c is then determined as the solution of the equation

$$\left(\frac{ds_1}{dx}\right)_{x=1} = \left(\frac{ds_2}{dx}\right)_{x=1},$$

which is, from (20) and (21),

$$z - 2\phi(\eta_1^{\mu+1}) = \{\phi(\eta_1^{\mu+1}) - \psi(\eta_1^{\mu+1})\} qV_1/kT, \quad (23)$$

where

$$\phi(x) = x \frac{d}{dx} \log \{F(x) \psi(x)\}. \quad (24)$$

When the interaction energy follows an inverse power law the quantities A_n , A'_n , etc., have been evaluated by Lennard-Jones and Ingham (1925). With their values the critical temperatures for $n = 4, 6, 10, 20$ and ∞ in the case of the simple cubic lattice are calculated from (23) and given below:

n	4	6	10	20	∞
kT_c/V_1	0.5778	0.8711	1.1342	1.2300	1.2331

In the case of the body-centred cubic lattice kT_c/V_1 is 0.944 for $n = 6$ and 1.738 for $n = \infty$. The case $n = \infty$ corresponds to Bethe's first approximation.

At temperatures below T_c the two curves s_1 and s_2 intersect at a second point with a non-vanishing s . The superlattice order s determined by this intersection point is calculated from the eqns. (20) and (21) and is shown as a function of T/T_c in fig. 2 for the simple cubic lattice and in fig. 3 for the body-centred cubic lattice. For the simple cubic lattice we have chosen the extreme values $n = 4$ and ∞ . Curves for intermediate n lie between these

two curves. For the body-centred cubic lattice, the values $n = 6$ and ∞ are chosen, because $n = 6$ gives the van der Waals attraction between neutral atoms and represents the interaction energy which falls off most slowly

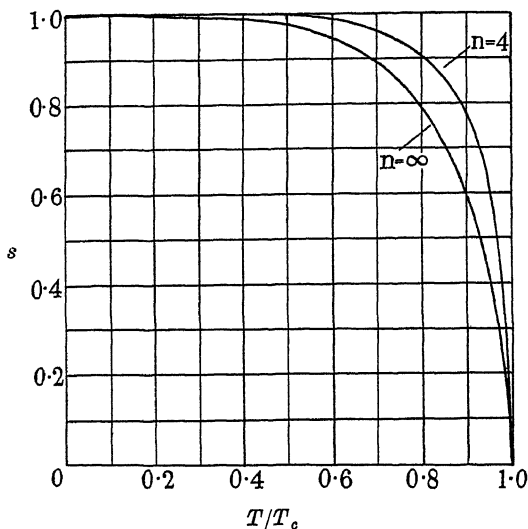


FIG. 2. The superlattice order s as a function of temperature for the simple cubic lattice. n is the index in the inverse power law assumed for V .

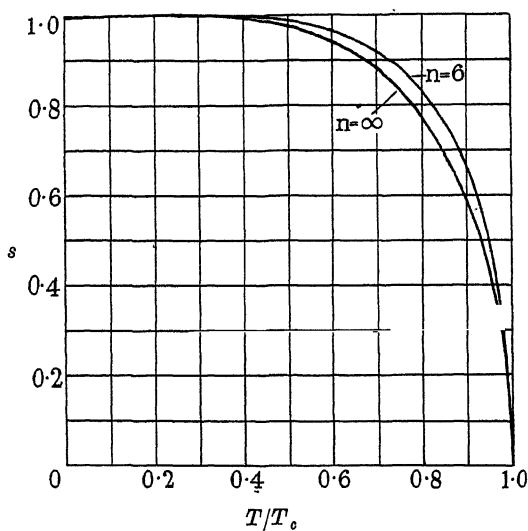


FIG. 3. The superlattice order s as a function of temperature for the body-centred cubic lattice. n is the index in the inverse power law assumed for V .

with distance when electrostatic forces are absent.† We shall discuss the law of interaction at the end of the next section.

It is seen from figs. 2 and 3 that the change of the superlattice order below the critical temperature is more sudden for small n than for large n . Since it has been found in the case of interaction only between nearest neighbours (Bethe 1935) that the change of the superlattice order just below T_c calculated by including the second shell is greater than that calculated by the first approximation, one would expect to increase this change still further by going into the second approximation on the present theory. In this way it might be possible to make the calculated configurational energy agree with experiment (cf. the next section), but the second approximation is too complicated for such a calculation to be practicable.

4. THE CONFIGURATIONAL ENERGY

With the help of (7), (10), (11) and the first equation of (12), the expression (26) in I for the configurational energy is reduced to

$$E/NV_1 = \frac{1}{4}\{(1+s)\psi_\eta + qs^2 + U^*/V_1\}, \quad (25)$$

where ψ_η is defined by (3), q by (5), and U^* by (42) in I. In particular, when the interaction energy follows an inverse power law, q is given by (6*a*) and (6*b*), and U^* by (46) and (50) in I for the simple cubic and the body-centred cubic lattice.

The variation of the configurational energy is very small above the critical temperature and the total change occurs practically within a very short interval just below the critical temperature. We therefore plot the expression $\{E(T_c) - E(T)\}/NkT_c$ against T/T_c for values of T from 0 to T_c and omit the insignificant part at T greater than T_c . The curves for the simple cubic lattice are shown in fig. 4 and those for the body-centred cubic lattice in fig. 5. The experimental curves‡ of Sykes (1935) for β brass (CuZn), which is a body-centred cubic crystal, are shown in fig. 5 for comparison. It can be seen that the change of the energy below the critical temperature is more sudden for small n than for large n . But the change is still not large enough when $n=6$ to account for the experimental data in the case of the body-centred cubic lattice.

Actually V cannot follow a simple inverse power law. It is known that atomic interactions can be adequately represented by two terms

$$\lambda r^{-m} - \mu r^{-n},$$

† See, for example, R. H. Fowler (1936).

‡ The curves for the energy are deduced from Sykes' data (Sykes 1935) on the specific heat by Bragg and Williams (1935, p. 559).

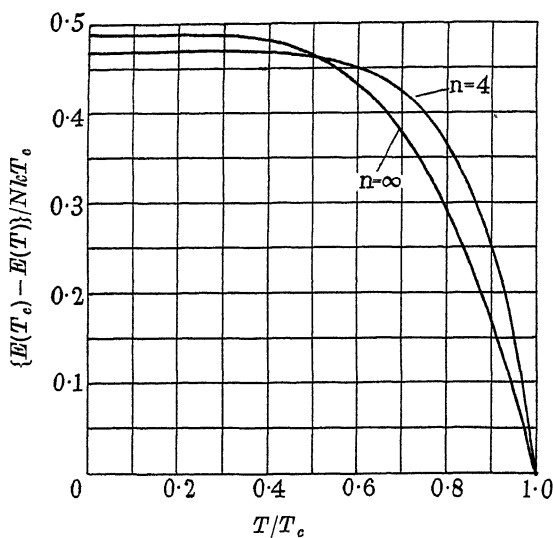


FIG. 4. The configurational energy as a function of temperature for the simple cubic lattice. n is the index in the inverse power law assumed for V .

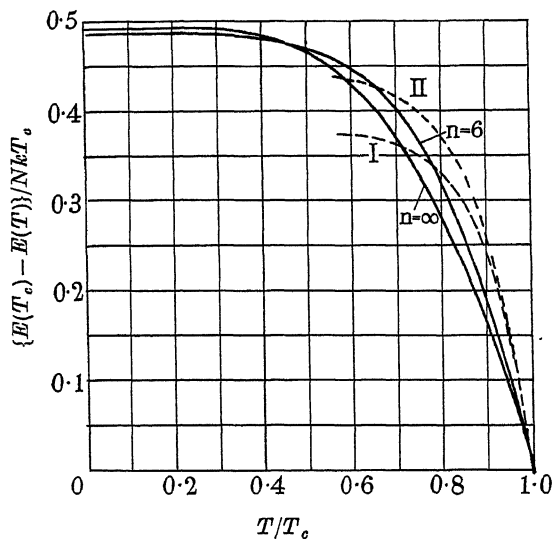


FIG. 5. The configurational energy as a function of temperature for the body-centred cubic lattice. n is the index in the inverse power law assumed for V . The two experimental curves are based on the two extreme assumptions as to the normal specific heat. The normal specific heat for the upper curve (II) is the mean for Cu and Zn, and the lower curve (I) is drawn by assuming Bethe's estimate of the abnormality in specific heat. See Bragg and Williams (1935, pp. 558-9).

where m is about 9 and n is 6. The first term arises from repulsion and the second term from the van der Waals attraction. Assuming this form of energy for each of V^{AA} , V^{BB} , V^{AB} we have

$$V^{AA} = \lambda^{AA}r^{-m} - \mu^{AA}r^{-n}, \text{ etc.},$$

and therefore

$$V = Ar^{-m} + Br^{-n},$$

where

$$A = \lambda^{AA} + \lambda^{BB} - 2\lambda^{AB}, \quad B = 2\mu^{AB} - \mu^{AA} - \mu^{BB}.$$

For the superlattice to be possible V must be positive, at least for $r = a_1$. This probably but not necessarily means that B is positive. But, if B were negative, A must be positive and V would fall off even more rapidly with distance than Ar^{-m} and would soon become negative. Such a V would evidently make the agreement between theory and experiment worse than a simple inverse 9th power law (since $m = 9$) and therefore much worse than the inverse 6th power law calculated above.

Now suppose B is positive. If A is also positive the result will be intermediate between the results obtained with each of the two terms separately and the agreement between theory and experiment will be worse than with a simple inverse 6th power law. If, on the other hand, A is negative, then as r increases V decreases more slowly than the inverse 6th power law given by the B term alone and the agreement with experiment will be improved.

Finally, I wish to thank Dr J. K. Roberts for his kind help.

SUMMARY

The general statistical theory of superlattices with long-range interaction given in the preceding paper is applied to the simple cubic lattice and the body-centred cubic lattice. It is shown that the assumption of an inverse power law for the interaction energy alters the calculated configurational energy in the direction required to make the agreement with experiment better than the assumption of interaction between nearest neighbours only, but the shift is not great enough to remove the discrepancy between theory and experiment completely.

REFERENCES

- Bethe, H. A. 1935 *Proc. Roy. Soc. A*, **150**, 552–75.
 Bragg, W. L. and Williams, E. J. 1935 *Proc. Roy. Soc. A*, **151**, 540–66.
 Fowler, R. H. 1936 “Statistical Mechanics”, 2nd ed., Chapter x. Camb. Univ. Press.
 Lennard-Jones, J. E. and Ingham, A. E. 1925 *Proc. Roy. Soc. A*, **107**, 636–53.
 Sykes, C. 1935 *Proc. Roy. Soc. A*, **148**, 422–46.
-

The nuclear spin of iodine

III. Further measurements upon the fine structures in the first spark spectrum

BY S. TOLANSKY, PH.D., AND G. O. FORESTER, B.SC.

Manchester University

(Communicated by P. M. S. Blackett, F.R.S.—Received 8 June 1938)

[Plate 1]

INTRODUCTION

In an earlier paper by Tolansky (1935*a*) measurements were given for the fine structures of forty-two lines in the visible region of the first spark spectrum of iodine. Of these eleven had then been classified by Murakawa (1933). It was proved that the nuclear spin of iodine is $\frac{5}{2}$. After these measurements had been communicated, Lacroute (1934) published an extensive analysis of the gross structure multiplets in the I^+ spectrum, the Zeeman effect being used as a guide. Lacroute established the positions of groups of terms going to the 4S , 2D and 2P series limits, the electron configurations involved being $5s^2 5p^4$, $5s^2 5p^3 6p$, $5s^2 5p^3 6s$, $5s^2 5p^3 5d$, $5s^2 5p^3 7s$, $5s^2 5p^3 6d$, $5s^2 5p^3 7s$. The eleven classified lines, whose fine structures had been analysed, were found to arise from transitions between terms going to the 4S series limit only. Fourteen of the remaining lines were classified by Lacroute who showed that they belong to the 2D system. A fine-structure analysis for these additional lines was then made by Tolansky (1935*b*). No intercombination lines between the two systems were available and the analysis proved to be difficult, with resulting ambiguity, in a number of the interval factors calculated. This we now know to be due to the existence of unexpectedly large interval factors in the $6p$ and $5d$ configurations (in what follows the $5s^2 5p^3$ group of electrons is omitted from the electron designations for the sake of brevity). A further complication exists due to the occurrence of perturbation.

Since the question of perturbations is of particular importance in connexion with the determination of a possible nuclear electrical quadrupole moment, it seemed desirable to extend the earlier measurements down into the ultra-violet region where strong intercombination lines appear. Since the earlier measurements were all made with a silvered Fabry-Perot inter-

ferometer, very accurate measurements could only be made up to the blue-green. To extend these into the ultra-violet, and to improve the previous measurements made in the blue and violet, we have made use of an exceptionally good quartz Lummer plate, a variable gap aluminized Fabry-Perot interferometer, and also a 21 ft. grating used up to the fourth order.

EXPERIMENTAL

As before, the I^+ spectrum was excited in a water-cooled hollow-cathode discharge tube through which helium and iodine were circulated. It is well known that the hollow-cathode discharge often selects out certain multiplets and considerably modifies their intensities relative to those given in spark and arc discharges. Most of the strongest lines excited in the Geissler tube and high-frequency discharge involve the $(^4S) 5d\ ^3D_2$ term, but in our hollow-cathode discharge these are almost suppressed. An examination of some early photographs taken with the 21 ft. grating showed that during the first few hours of life of the tube, these lines were reasonably strong, but on ageing they disappeared (see note at end of paper).

The structures of the lines in the violet and ultra-violet were mainly studied with the aid of a Hilger quartz Lummer plate, 20 cm. long and 0.342 cm. thick. This was housed in a constant-temperature room. It was crossed either with a two-prism glass spectrograph or a quartz spectrograph. When using the glass instrument, the interferometer was placed in the parallel beam just before the prisms, and as the camera lens has a focal length of 125 cm. a very large dispersion results. A thin Lummer plate was specially chosen in order to obtain a wide range of measurement. By means of a cross-slit, the extraordinary image could be completely separated from the ordinary image. Thus the best resolution is obtained without auxiliaries by simply making use of the double refraction of the interferometer itself.

Since surface scratches and dust cause considerable scatter in the ultra-violet, the Lummer plate was placed before the slit when used with the quartz spectrograph, the fringes being projected on to the slit with a quartz fluorite achromat. The focal length is such that the image upon the plate was equal in size to that which would be given by a 75 cm. focal length camera lens if the interferometer had been placed in the parallel beam.

To solve difficulties due to overlapping of orders when very widespread structures were encountered, the spectrum was also photographed with a large-aperture variable-gap aluminized Fabry-Perot interferometer. The resolving power of this instrument was considerably less than that of the Lummer plate, but sufficed to settle problems of overlapping. The resolving

power of the Lummer plate is very high, and it is in fact possible to separate lines only half as far apart as that permitted by the Rayleigh definition of resolving power. Thus in the green an effective resolving power of some 600,000 could be attained. Even in this region, where good reflecting coefficients are available, a thin Lummer plate is superior to a Fabry-Perot interferometer with equivalent order separation. For example, at 5500 Å the order separation with our Lummer plate is 1.2 cm.^{-1} , which is that produced by a Fabry-Perot with a gap of 4.16 mm. Such a Fabry-Perot has, in this region, a resolving power certainly no greater than 300,000. Clearly when the width of a complex pattern is such that the maximum Fabry-Perot gap available is less than 5 mm., then the Lummer plate is superior to the Fabry-Perot, even in the red. Since there are many wide patterns, we have been able to improve upon some of the earlier reported measurements as well as make new observations.

The spectra were photographed with Ilford and Eastman plates, exposure times varying from 15 sec. to 4 hr. for the interferometers, and 12 hr. for the grating. Plate 1 shows an enlargement of a part of the spectrum taken with the Lummer plate.

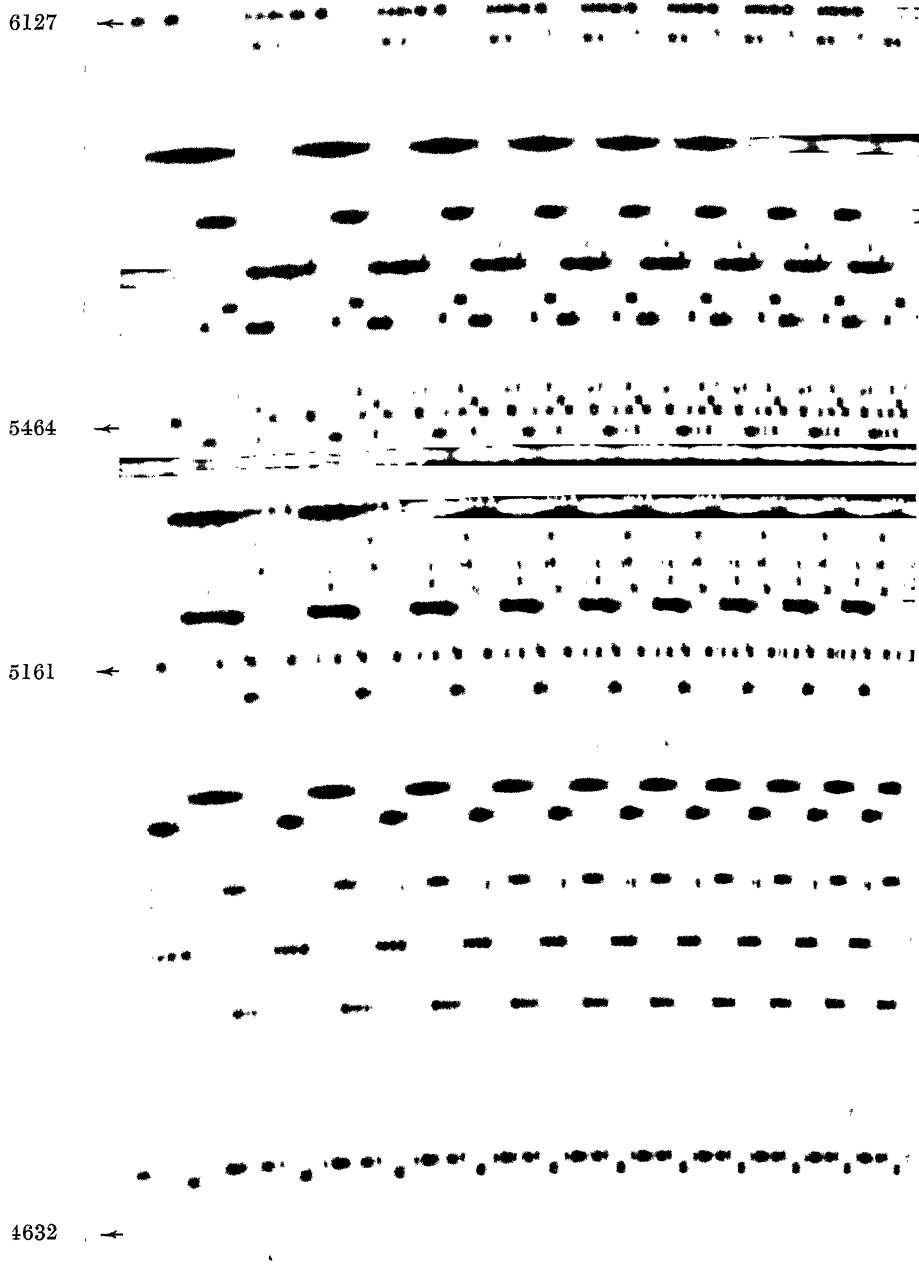
OBSERVATIONS

Table I contains the lines whose structures have so far been accurately measured (see note at end of paper). Those lines marked with an asterisk are new measurements, the rest being taken from Part I. The units used are $\text{cm.} \times 10^{-3}$. The visual estimate for the intensity of each component is indicated underneath in brackets. The classifications in column 2 are those given by Lacroute, excepting where these have been modified because of existing errors in Lacroute's paper (these will be mentioned later). The lines are classed A, B, C, according to whether the resolution in the pattern is good (A), moderate (B) or bad (C). In Part I this was affected by the falling off in resolving power of the silvered Fabry-Perot interferometer with decreasing wave-length. In the present list, the complexity or narrowness of the patterns are the factors affecting the class.

ANALYSIS

A. The ^4S system

Of the fifty lines in Table I only thirty-one have as yet been classified, the transitions for these being shown in fig. 1. There are two main self-contained groups, the ^4S and ^2D groups of terms, these being connected by a small number of intercombination lines. The latter are relatively infrequent in



Fine structure in the I^+ spectrum.

(Facing p. 80)

TABLE I. FINE STRUCTURE IN THE IODINE SPARK LINES

Wave-length	Classification	Structure $\text{cm.}^{-1} \times 10^{-3}$								Class
		Red				Violet				
6585.0	(² D) 5d ³ D ₁ -(² D) 6p ³ D ₁	0	47	178	256					A
		(2)	(6)	(3)	(2)					
6516.1		0	153	327	502	628	779	834	926	A
		(1)	(3)	(12)	(2)	(4)	(5)	(4)	(4)	(3)
6204.7		0	71	164	244	393				A
		(3)	(5)	(1)	(2)	(2)				
6161.9		0	579	949						A
		(4)	(3)	(2)						
*6127.4	(² D) 6s ³ D ₂ -(² D) 6p ³ D ₁	0	215	244	360	433	535	636	650	A
		(8)	(6)	(2)	(2)	(4)	(3)	(3)	(2)	
6068.8	(² D) 5d ³ D ₁ -(² D) 6p ³ F ₂	0	35	126	251	308				A
		(2)	(3)	(10)	(4)	(15)				
5950.1	(⁴ S) 6s ³ S ₁ -(⁴ S) 6p ³ P ₁	0	47	121	204					C
		(3)	(3)	(2)	(4)					
5920.7	(⁴ S) 6p ³ P ₀ -(⁴ S) 7s ³ S ₁	Single								B
5890		0	57	150						C
		(4)	(3)	(2)						
5813		Single								A
5787.1	(⁴ S) 6p ³ P ₂ -(⁴ S) 7s ³ S ₁	0	56	106	149					B
		(4)	(1)	(2)	(2)					
5774.7	(⁴ S) 5d ³ D ₂ -(⁴ S) 6p ⁵ P ₁	0	741	1291	1686	1913				A
		(5)	(4)	(3)	(2)	(1)				
5760.4	(² D) 5d ³ D ₁ -(² D) 6p ³ D ₂	0	56	100	132	159				B
		(10)	(3)	(5)	(4)	(4)				
5678.0	(² D) 6s ³ D ₂ -(² D) 6p ³ F ₂	0	86	173	238	293	424	482		A
		(1)	(3)	(10)	(4)	(3)	(3)	(6)		
5625.7	(⁴ S) 6s ³ S ₁ -(⁴ S) 6p ³ P ₂	0	28	70						C
		(2)	(2)	(3)						
5598.5		0	395	618	813	1037				A
		(5)	(18)	(1)	(3)	(9)				
5593.1		0	26	80	121	151				B
		(7)	(10)	(2)	(4)	(4)				
5504.8	(⁴ S) 6s ³ S ₁ -(⁴ S) 6p ³ P ₀	0	33							C
		(1)	(1)							
5496.8	(⁴ S) 6s ⁵ S ₂ -(⁴ S) 6p ⁵ P ₁	0	429	756	991	1141				A
		(5)	(4)	(3)	(2)	(1)				
*5464.8	(⁴ S) 6s ⁵ S ₂ -(⁴ S) 6p ⁵ P ₂	0	36	395	449	659	705	866	893	A
		(1)	(18)	(12)	(4)	(5)	(6)	(4)	(4)	(4)
*5407.3	(² D) 6s ³ D ₂ -(² D) 6p ³ D ₂	0	197	234	372	415	519	559	644	A
		(8)	(6)	(4)	(4)	(3)	(2)	(3)	(2)	
5345.2		0	151	305	419	507	536	605		A
		(1)	(12)	(10)	(8)	(6)	(4)	(2)		

TABLE I (continued)

Wave-length	Classification	Structure $\text{cm.}^{-1} \times 10^{-3}$						Class
		Red						
5338.2	$(^2\text{D}) 6s^3\text{D}_2 - (^2\text{D}) 6p^3\text{D}_3$	0	107	236	348	429	461	B
		(1)	(8)	(6)	(4)	(2)	(1)	
5269.4		0	101	216	353	490	623	B
		(8)	(10)	(1)	(6)	(12)	(8)	
5234.6		0	192	355	502	622	701	A
		(7)	(6)	(5)	(4)	(3)	(2)	
5228.9		0	39	84	176			C
		(1)	(3)	(2)	(4)			
5216.2	$(^2\text{D}) 6s^3\text{D}_1 - (^2\text{D}) 6p^3\text{F}_2$	0	22	79	122	199	283	B
		(3)	(1)	(3)	(4)	(5)	(6)	
5214.0		0	36	83				C
		(2)	(3)	(3)				
*5161.2	$(^4\text{S}) 6s^5\text{S}_2 - (^4\text{S}) 6p^5\text{P}_3$	0	396	696	907	1034		A
		(5)	(4)	(3)	(2)	(1)		
5156.4	$(^4\text{S}) 6p^5\text{P}_1 - (^4\text{S}) 7s^5\text{S}_2$	0	75	163	266	382		A
		(2)	(3)	(4)	(5)	(6)		
5125		0	70	118	182			C
		(2)	(2)	(5)	(1)			
4987.0	$(^2\text{D}) 6s^3\text{D}_1 - (^2\text{D}) 6p^3\text{D}_2$	0	38					C
		(4)	(3)					
4806.4	$(^2\text{D}) 5d^3\text{F}_2 - (^2\text{D}) 6p^3\text{F}_2$	0	53	121	168	213	252	B
		(4)	(5)	(5)	(6)	(6)	(12)	
*4675.5	$(^2\text{D}) 5d^1\text{D}_2 - (^2\text{D}) 6p^1\text{D}_2$	0	59	253	293	383	539	A
		(8)	(2)	(2)	(6)	(12)	(3)	
4666.5		Single						A
4632.4	$(^4\text{S}) 6s^5\text{S}_2 - (^4\text{S}) 6p^3\text{P}_2$	0	399	695	928	1054		A
		(5)	(4)	(3)	(2)	(1)		
4561.0	$(^2\text{D}) 5d^3\text{F}_2 - (^2\text{D}) 6p^3\text{D}_3$	0.....	228					B
		(1)			(4)			
4544.3	$(^2\text{D}) 5d^3\text{D}_2 - (^2\text{D}) 6p^1\text{D}_2$	0	240	373	685			C
		(3)	(2)	(5)	(1)			
4540.9	$(^2\text{D}) 6s^3\text{D}_2 - (^2\text{D}) 6p^3\text{P}_2$	0	110	162				C
		(2)	(1)	(1)				
4488.5	$(^4\text{S}) 5d^5\text{D}_1 - (^2\text{D}) 6p^3\text{D}_1$	0	117	162				B
		(2)	(1)	(1)				
4456.9		0	50					C
		(5)	(4)					
*4403		0	122	310	566	892		A
		(1)	(2)	(3)	(4)	(5)		

TABLE I (continued)

Wave-length	Classification	Structure $\text{cm.}^{-1} \times 10^{-3}$					Class
		<div>RedViolet</div>					
*4225.5	$(^2\text{D})\ 5d\ ^3\text{F}_2-(^2\text{D})\ 6p\ ^1\text{F}_3$	0	154	370	646	993	B
		(5)	(9)	(8)	(10)	(14)	
*4129		0	116	215	262	304	B
		(5)	(4)	(3)	(2)	(1)	
*4036.1	$(^2\text{D})\ 5d\ ^3\text{F}_2-(^2\text{D})\ 6p\ ^3\text{F}_3$	0	138	319	543	806	A
		(2)	(3)	(4)	(5)	(6)	
*3892.9		0	120	264			B
		(3)	(4)	(5)			
*3833.7		0	160	352	653	1051	A
		(1)	(2)	(3)	(4)	(5)	
*3779.4	$(^4\text{S})\ 5d\ ^5\text{D}_3-(^2\text{D})\ 6p\ ^3\text{D}_2$	0	73	201	365	557	A
		(2)	(3)	(4)	(5)	(6)	
*3302.4	$(^4\text{S})\ 5d\ ^3\text{D}_2-(^2\text{D})\ 6p\ ^3\text{D}_2$	0	677	1220	1556	1806	A
		(5)	(4)	(3)	(2)	(1)	
*3209.6	$(^4\text{S})\ 6s\ ^5\text{S}_2-(^2\text{D})\ 6p\ ^3\text{D}_2$	0	378	663	866	965	A
		(5)	(4)	(3)	(2)	(1)	

See note at end of paper.

the spectrum. In the earlier reports no intercombination lines had been measured, and as a result only approximate estimates could be given for the fine-structure term interval factors in the ^2D system, since the structures of lines in this group are difficult to analyse. Fortunately, a number of the ^4S terms have very large interval factors leading to considerable simplification of many patterns. We shall therefore first consider lines belonging to the ^4S system.

The analysis of the ^4S lines has already been given in Part I, but since more lines and higher resolving power are now available, the improved data lead to a better and more self consistent analysis. In Part I it was necessary to employ the Fisher-Goudsmit geometrical method of analysis with its attendant uncertainties. This can now be dispensed with and an analysis by the usual method of frequency differences has been carried through. The final analysis shows that the general trend of the previous results is correct especially with the larger interval factors. The previous results are in error to the order of about 10 %, apart from the very small values which are most easily affected by the graphical method.

It was shown in Part I that the simplicity of most of the line patterns based upon the $(^4S) 6s\ ^5S_2$ and $(^4S) 5d\ ^3D_2$ terms is due to the fact that the structures in these two terms are relatively wide. The three lines $(^4S) 6s\ ^5S_2 - (^4S) 6p\ ^5P_1\ ^5P_3\ ^3P_2$ (5497, 5161, 4632) each exhibit a simple regular quintet pattern degrading to the violet and extending respectively over $1141, 1025, 1054\text{ cm.}^{-1} \times 10^{-3}$. From this it is clear that the approximate width of the

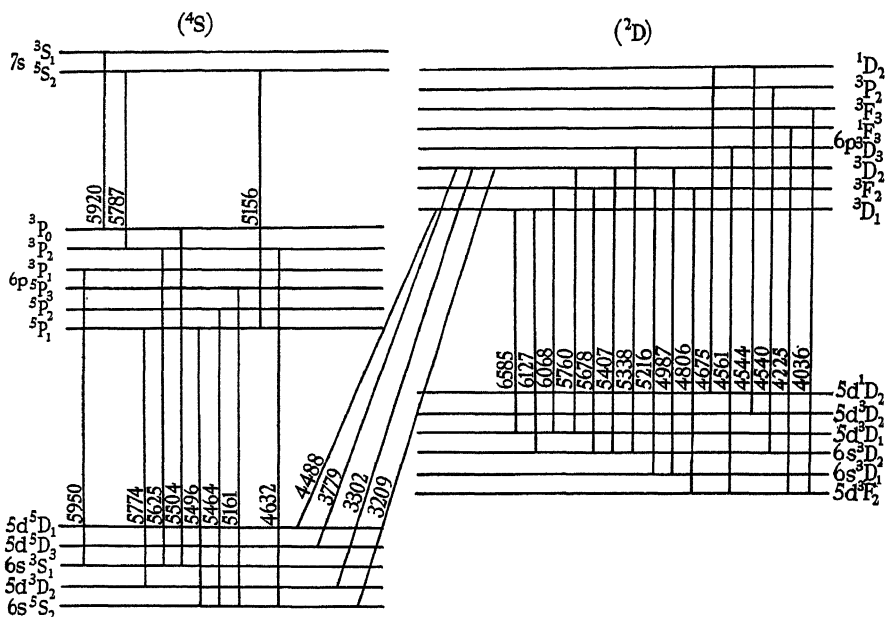


FIG. 1

structure in the $(^4S) 6s\ ^5S_2$ term is of the order of $1050\text{ cm.}^{-1} \times 10^{-3}$. Since the small structure due to the upper term is not resolvable in these lines an accurate analysis could not previously be carried out without the aid of some arbitrary assumption based upon the electron coupling constants for the upper terms. It was in fact not possible to decide for instance whether the width of the structure in the term was 1025 or $1141\text{ cm.}^{-1} \times 10^{-3}$. If the former, the latter value arises from a negative upper interval factor. If 1141 be correct then the 1025 interval implies a positive upper-interval factor for its line transition. This is the source of ambiguity in Part I.

With the aid of the Lummer plate we have been able to achieve sufficient resolution of the upper-term structure for the line $(^4S) 6s\ ^5S_2 - (^4S) 6p\ ^5P_1$ (5464) to permit of a complete analysis. The detailed analysis is shown in fig. 2. Complete resolution of all the close components is not attained and is in fact hardly possible. It is practically certain that the very small differ-

ences between the calculated and measured patterns arise from this only. In all the diagrams the units are $\text{cm.}^{-1} \times 10^{-3}$.

The $(^4\text{S})\ 6s\ ^5\text{S}_2$ term is spherically symmetrical and therefore not perturbed by any existing nuclear electrical quadrupole moment. Hence the

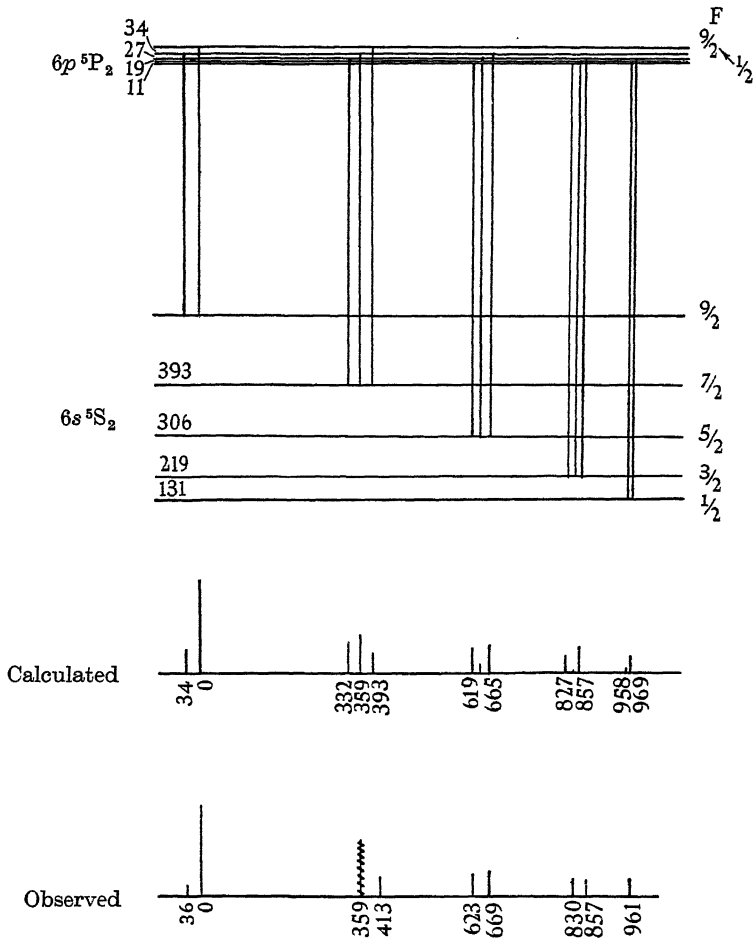


FIG. 2. $\lambda\ 5464$.

interval rule should be strictly obeyed. This we find to be true within the limits of our measurements. Defining the interval factor as A , where the separation between two fine structure levels $F+1$ and F is $A(F+1)$, we get $A = 87.4\ \text{cm.}^{-1} \times 10^{-3}$ for the term in question. This is 12 % less than that given in Part I.

The analysis for 5161 is shown in fig. 3. This line exhibits very sharp individual components (see Plate 1), and as expected from this the upper-term structure turns out to be very narrow indeed, extending only over about $7\frac{1}{2} \text{ cm.}^{-1} \times 10^{-3}$. The calculated and observed separations are in

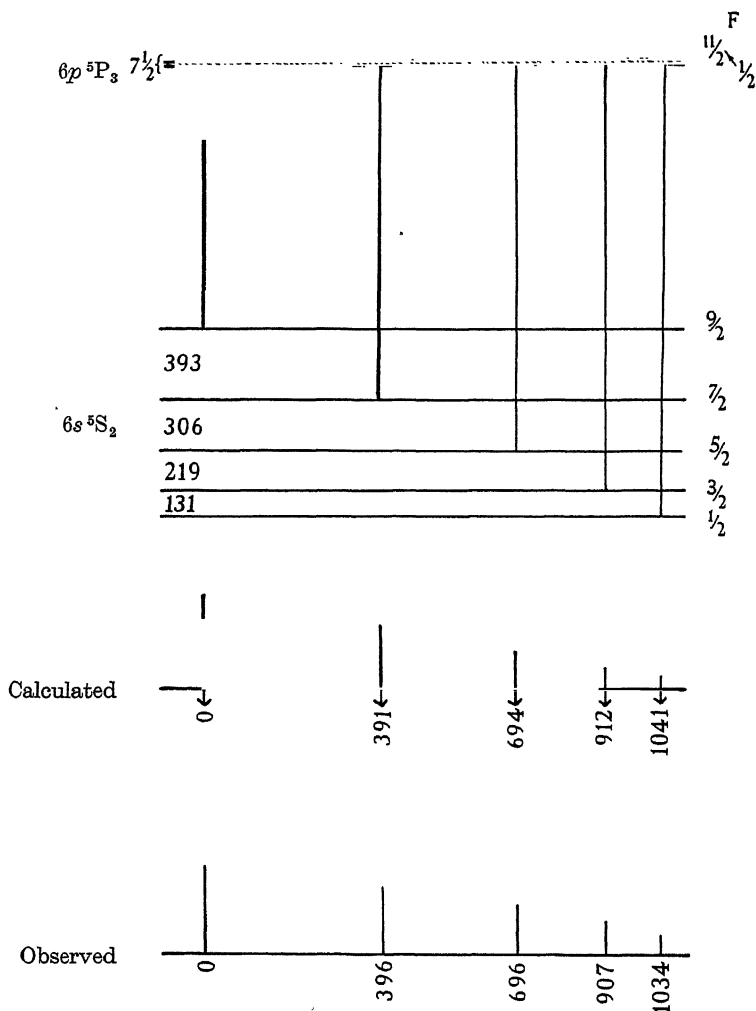


FIG. 3. λ 5161.

agreement to within $3 \text{ cm.}^{-1} \times 10^{-3}$, apart from the first pair in which there is a difference of $5 \text{ cm.}^{-1} \times 10^{-3}$. This is still very small, and is occasioned by the fact that the first two components are inherently broader than the others.

As shown in fig. 4 the pattern for $(^4S) 6s \ ^5S_2 - (^4S) 6p \ ^5P_1$ (5497) indicates

that the upper term has in this case an inverted structure. In this, as in later diagrams, the arrow beneath a group of lines indicates the position of the calculated optical centre of gravity for the unresolved group. The

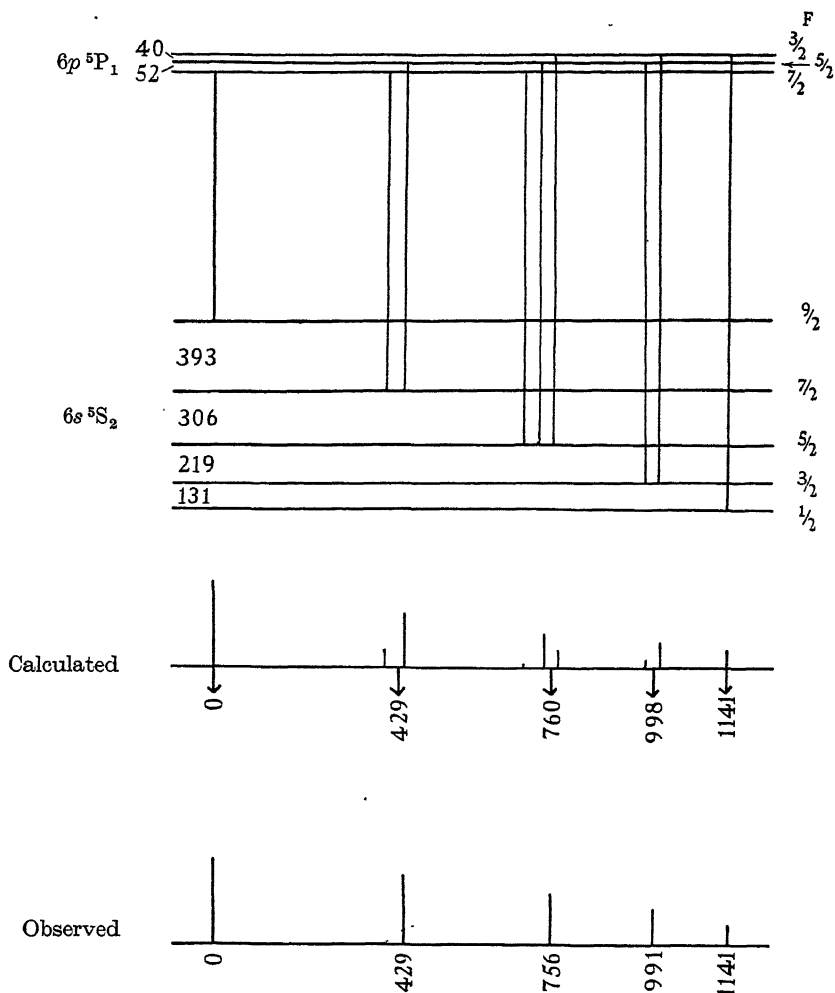


FIG. 4. $\lambda\ 5497$.

positions of these agree with those of the observed components. The apparent deviation from the interval rule for the upper term has no significance. The structure just derived for the upper term can now be applied to the line $(^4S)\ 5d\ ^3D_2 - (^4S)\ 6p\ ^5P_1$ (5774), the analysis of which is shown in fig. 5.

The line transitions to $(^4S)\ 5d\ ^3D_2$, which are the strongest in Geissler tube and high-frequency discharges, are peculiarly weakened in the hollow

cathode. We have only succeeded in measuring two of the many lines with the interferometers. A number of others were found upon a plate taken with the 21 ft. grating. These are discussed in a note at the end of this paper.

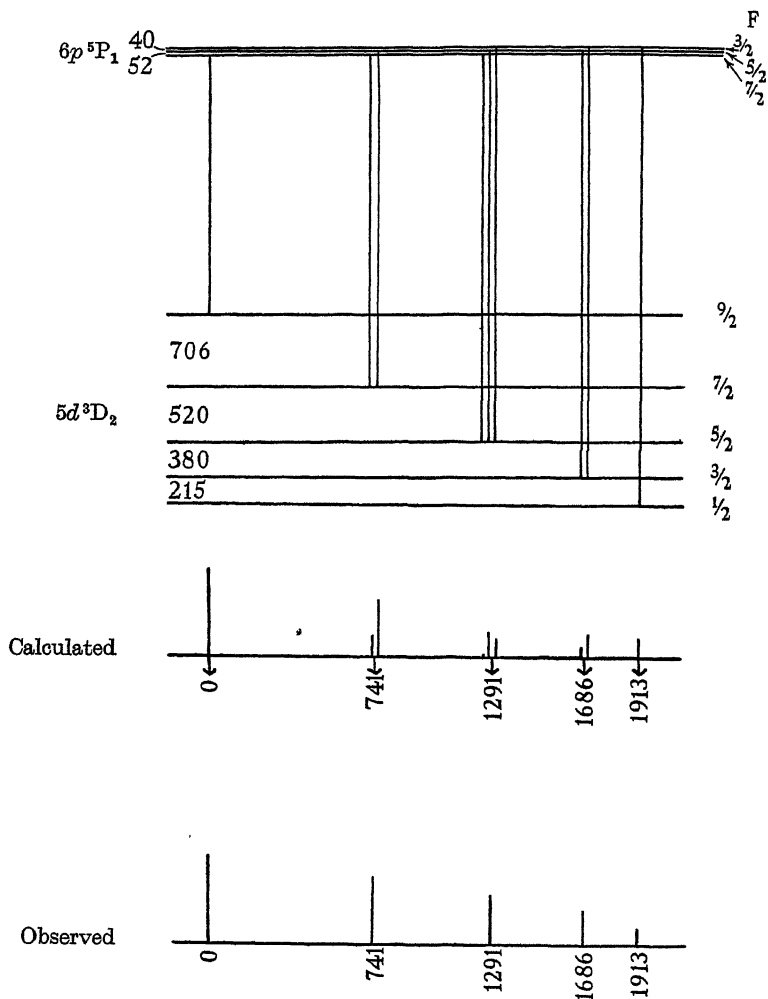


FIG. 5. λ 5774.

The ultra-violet lines $(^4S) 6s^5S_2-(^2D) 6p^3D_2$ (3209) and $(^4S) 5d^3D_2-(^2D) 6p^3D_2$ (3302) are of particular value, since they are intercombinations between the two systems and in both the structure of the lower term is well established. The value derived for the upper-term structure therefore constitutes an important link between the (^4S) and (^2D) systems. In both lines the individual components are broadened by the existence of the non-

resolved upper structure, yet although the accuracy is limited a satisfactory analysis can be carried out (see fig. 6). The structure in the line 3302 is so wide that the pattern is overlapped in the Lummer plate photograph, and as the resolving power of a small-gap aluminium Fabry-Perot is not very high there is an appreciable error in the estimated positions of two of the components.

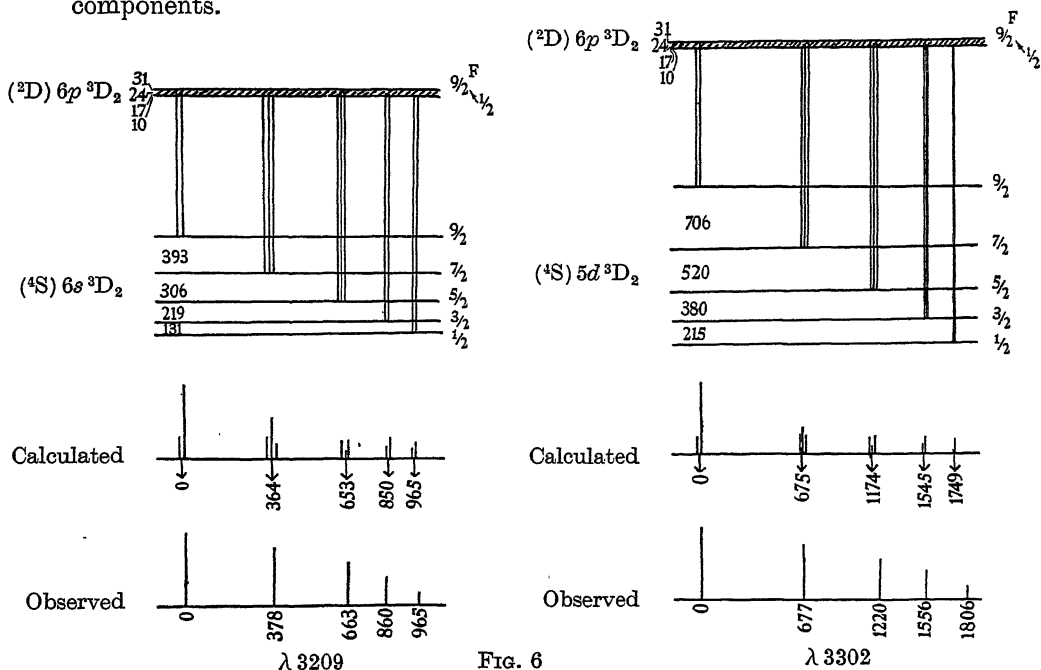


FIG. 6

The calculated and observed patterns for $(^4S) 6s \ ^5S_2 - (^4S) 6p \ ^3P_2$ (4632) are shown in fig. 7a. The upper-term structure is small, extending only over $14 \text{ cm}^{-1} \times 10^{-3}$. It is seen from the plate that the individual components of this line are quite sharp. The data from this analysis can now be applied to 5625 which gives the structure of $(^4S) 6s \ ^3S_1$ (see fig. 7b).

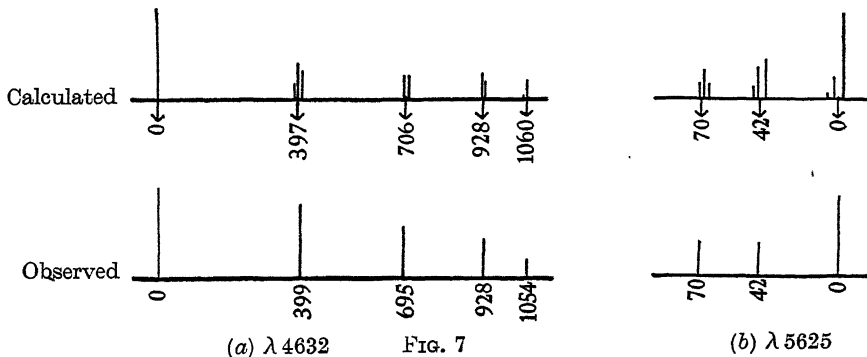


FIG. 7

The above method of analysis can be extended to all the lines in the (4S) system. We failed to observe the expected weak sixth component in 3779. It was not found possible to fit 5787 into the analysis scheme, and we suggest that this line has been incorrectly allocated. The list of interval factors given in Table II is much more reliable than the corresponding list in Part I, since we have succeeded in resolving and analysing some upper-term structures.

B. The (2D) system

The difficulties previously encountered with the (2D) terms arose from the existence of perturbation. We have been able to analyse the (2D) structures without the aid of the (4S) terms, and the essential correctness of the whole interpretation is shown by the fact that the intercombination lines fit the scheme, the structure derived for the common (2D) $6p\ ^3D_2$ being the same when the two systems are used independently.

The line (2D) $5d\ ^3D_1$ -(2D) $6p\ ^3D_1$ (6585) has few components, and as shown in fig. 8*a* the resulting narrow pattern can be fitted into an analysis. From this the structure derived for the upper term is small. This can now be applied to the elucidation of (2D) $6s\ ^3D_2$ -(2D) $6p\ ^3D_1$ (6127). Even if the previously derived value for the upper term is entirely disregarded, it is impossible to fit this pattern into a scheme based upon the interval rule. It

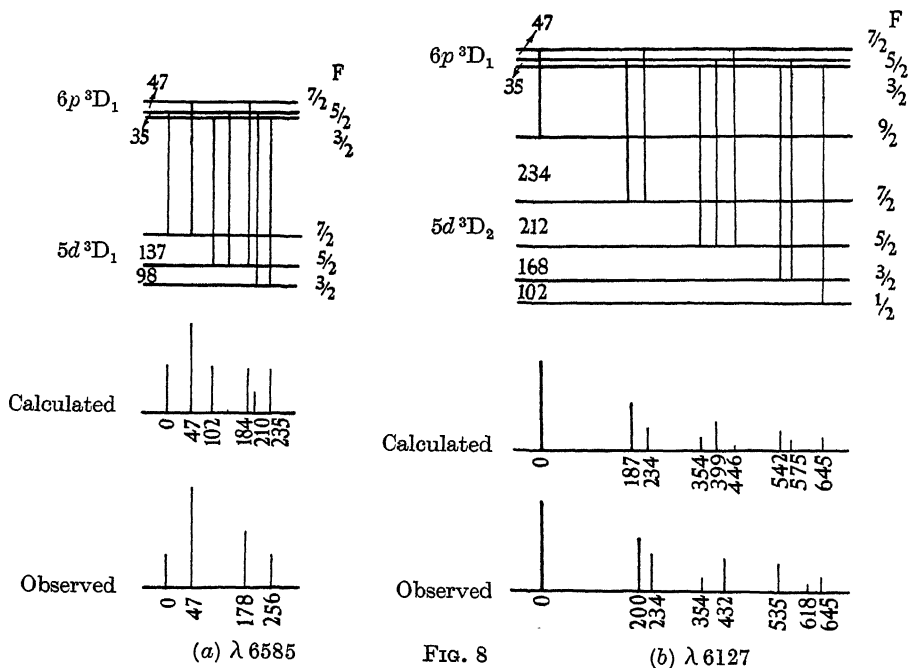


FIG. 8

is clear that the lower term is perturbed. As shown in fig. 8*b* the observed and calculated patterns fit fairly closely.

The analysis of the line (^2D) $6s\ ^3\text{D}_2$ -(^2D) $6p\ ^3\text{D}_2$ (5407) completely confirms the conclusion about perturbation. This line has been very carefully measured with the Lummer plate and eight components could be clearly distinguished. It is difficult to obtain very exact measurements of the closely grouped pairs, but we have been able to measure the positions of components 374 and 415 with high accuracy. The analysis is altogether much more reliable than that of the two previously considered (^2D) lines. It is shown in fig. 9. Two important points emerge. First, the lower term is

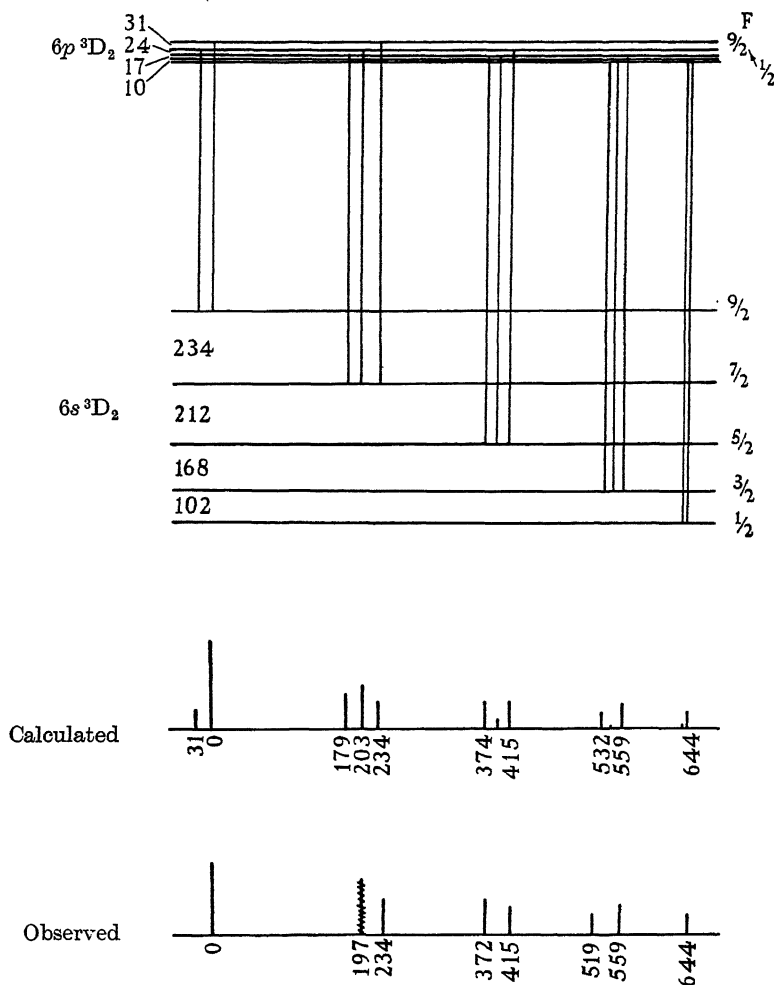


FIG. 9. λ 5407.

quite definitely perturbed, secondly, the structure derived independently for the upper $(^2D) 6p^3D_2$ term is in strikingly exact agreement with that found for the same time by means of the (^4S) intercombination lines.

A comparison is made in fig. 10 between the calculated and observed patterns for some of the remaining lines, the agreement being satisfactory. Two lines, $(^2D) 5d^3F_2-(^2D) 6p^1F_3$ (5225) and $(^2D) 5d^3F_2-(^2D) 6p^3F_3$ (4036) only appeared upon one photograph taken with the Lummer plate and failed to reappear as the tube aged. The lines are weak and the structures wide. A complete order is filled in the first case so that the predicted weak sixth component falls upon the first component of the next order. In the second case the line was not strong enough for the last weak component to be observed.

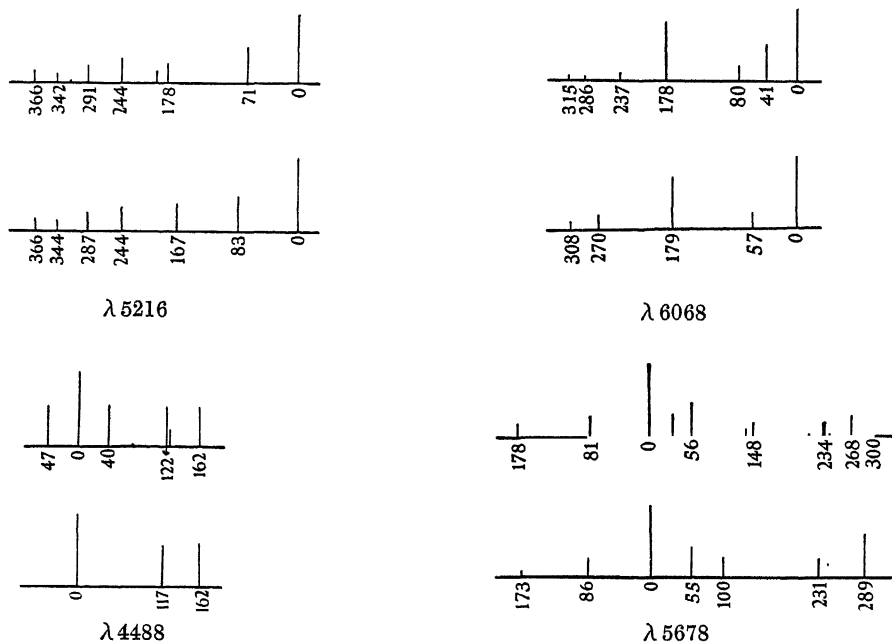


FIG. 10

The lines $(^2D) 5d^1D_2-(^2D) 6p^1D_2$ (4675) and $(^2D) 5d^3D_2-(^2D) 6p^1D_2$ (4544) require special mention, since the three terms involved are not linked to any others. The structures are almost identical, both exhibiting highly complex patterns indicating thereby large structures in upper and lower terms. The agreement between the observed and calculated structures is reasonably good. However, an alternative analysis is possible if it is assumed that all the three terms have inverted structures. It seems very improbable that this is so, but we have evaluated these alternative structures and included

them as a footnote in Table III. Three other lines only need be mentioned. The fit in 4806 is very approximate, and it is possible that some confusion has been introduced by a line 2 Å away. Poor resolution has been achieved for 4540, since it only appeared upon old Fabry-Perot photographs. It had been previously reported that the strongest component in 5216 had a close very weak neighbour (intensity only one-twenty-fourth that of the strong line). This does not appear to be genuine.

TERM INTERVAL FACTORS FOR THE TWO SYSTEMS

When the Landé interval rule is strictly obeyed, the interaction energy between the nuclear spin I and the total angular momentum of the electron system J is governed by the law $E = A \cdot IJ \cdot \cos(IJ)$. The interval factor, A , is a measure of the coupling between nuclear and external magnetic moments. It follows that the interval between two fine structure levels with

TABLE II. FINE STRUCTURE INTERVAL FACTORS IN THE
(⁴S) SYSTEM OF TERMS, CM.⁻¹ × 10⁻³

$5s^2 5p^3 6s$	$5s^2 5p^3 7s$	$5s^2 5p^3 6p$	$5s^2 5p^3 5d$
⁵ S ₂ 87.4	⁵ S ₂ 24.4	⁵ P ₁ -15.3	⁵ D ₁ 35
		⁵ P ₂ 7.5	⁵ D ₃ -30.3
		⁵ P ₃ 0.9	
³ S ₁ -9	³ S ₁ ~0	³ P ₁ 28	³ D ₂ 152*
		³ P ₂ -1.1	
		³ P ₀ 0	

* Perturbed.

TABLE III. FINE STRUCTURE INTERVAL FACTORS IN THE
(²D) SYSTEM OF TERMS, CM.⁻¹ × 10⁻³

$5s^2 5p^3 6s$	$5s^2 5p^3 6p$	$5s^2 5p^3 5d$
³ D ₁ 20.3	³ D ₁ 13.6	³ F ₂ 9.3
³ D ₂ 58.4*	³ D ₂ 6.8	
	³ D ₃ 19.4	
	³ F ₂ 39.6	³ D ₁ 39.1
	³ F ₃ ~53	³ D ₂ 69 (a)†
	¹ F ₃ 68.7	¹ D ₂ 69.8 (b)†
	³ P ₁ ~37‡	
	³ P ₂ ~45	
	¹ D ₂ 89.8 (c)†	

* Perturbed.

† Alternatively: (a) = -89, (b) = -89.8, (c) = -69.8.

‡ See note at end of paper.

quantum numbers $F + 1$ and F is $A(F + 1)$. In Tables II and III are shown the observed interval factors for the (4S) and (2D) terms respectively, in units of $\text{cm.}^{-1} \times 10^{-3}$. The perturbed structures, indicated by an asterisk, will be discussed later.

DISCUSSION

The data given in Tables II and III are much more complete and more accurate than those given earlier. Considering first Table II, it is clear that a number of anomalies exist. The interval factor for $6s\ ^5S_2$ is 3.6 times that for $7s\ ^5S_2$. Both terms are formed by the addition in parallel of the s electron to the electron group forming the basic ion term $5p^3\ ^4S_{\frac{3}{2}}$ of I^{++} . Clearly this can be explained completely by assuming the nuclear coupling of the $5p^3$ group to be small, since one would expect a $7s$ electron to give a coupling constant no greater than one-fourth that of a $6s$ electron. However, the 3S_1 interval factors fail to fit this scheme. They arise from the addition of the s electron anti-parallel to the $5p^3$ group of the $^4S_{\frac{3}{2}}$ term of I^{++} . According to the vector coupling theory of White (White 1930) the interaction energy is proportional to $\cos JS$, which in this case equals -1 . This means a negative interval factor, numerically smaller than that of the positive factor for the 5S_2 term. The interval factor found for $6s\ ^5S_2$ is 87.4, whilst that for $6s\ ^3S_1$ is -9 . This therefore fits theory. However, the smallness of the $6s\ ^3S_1$ factor can only be accounted for by assuming that the coupling constant for the $5p^3$ electrons is approximately the same as that of the $6s$ electron. This contradicts the previous conclusion about the $6s\ ^5S_2$ and $7s\ ^5S_2$ terms. Furthermore, if the $5p^3$ coupling does approximate that of $6s$ then it will be clearly much greater than that of the $7s$ electron, from which it follows that $7s\ ^3S_1$ should have a larger interval factor than $6s\ ^3S_1$ and, in fact, one that is approximately half that of $6s\ ^5S_2$.

It follows that if the coupling constant of the $5p^3$ group is large, the interval factor for $7s\ ^3S_1$ is too small. Conversely, if the $5p^3$ coupling factor is small, then the observed interval factor for $6s\ ^3S_1$ is too small. The remaining data show that the $5p^3$ group has a considerable degree of coupling with the nucleus, hence it must be concluded that the $7s\ ^3S_1$ interval factor is too small.

The third column shows clearly that the coupling constant for the $6p$ electron is appreciable. Since theory shows that in jj coupling a $p_{\frac{1}{2}}$ electron has semi-penetrating properties, it is not surprising that the $5p_{\frac{1}{2}}$ electrons in the $5p^3$ group result in producing a constant for the latter comparable with that of a $6s$ electron.

If we disregard the anomalously large $5d^3D_2$, the figures in the fourth column can be accounted for by attributing a coupling constant of the order of 33 to the $5p^3$ group and 2 to the $5d$ electron. These values are in reasonable agreement with the rest of the data. The derivation is crude, but suffices to show that the coupling constant for the $5p^3$ group is large enough to explain the data for the other three columns (apart from $7s^3S_1$).

The strikingly large interval factor for $5d^3D_2$ deserves particular mention. Attention has already been drawn to this in the earlier report. The present measurements show that not only is the structure exceptionally large but is also perturbed, failing to obey the interval rule (this will be discussed in detail later). It is impossible to attribute penetrating properties to a $5d$ electron, even in jj coupling. As pointed out in a note, at the end, added in proof, Lacroute's classification is incorrect. This term belongs to an sp^5 configuration.

We shall consider now the (2D) terms in Table III, the interpretation of which is more difficult. The wide difference between the values of the interval factors of the $6s^3D_2$ and $6s^3D_1$ terms in the first column also proves that the $5p^3$ core group has a coupling constant in the (2D) system comparable with that of the $6s$ electron. Applying this conclusion to the $6p$ terms in the second column, we find the following. The series limit (2D) having $L = 2$ may possibly lead to a higher $5p^3$ coupling constant than the series limit (4S), but even allowing for this it seems quite certain that the interval factors for $6p^1F_3$ (68.7) and $6p^1D_2$ (89.8) are much greater than one would expect. These terms are based upon the same series limit term ($^2D_{\frac{3}{2}}$, since the multiplets are normal) as, for instance, $6p^3D_3$ for which the interval factor is 19.4. Coupling theory shows that the series limit configuration contains only one $p_{\frac{1}{2}}$ electron. Clearly, then, the structure in $6p^1D_2$ is too great and the same is almost certainly true for $6p^1F_3$.

Finally, from the third column we find as follows. The $5d^3F_2$, $5d^3D_1$ and $5d^3D_2$ terms go to one series limit ($^2D_{\frac{3}{2}}$) and the remaining $5d^1D_2$ goes to the $^2D_{\frac{3}{2}}$ limit. Since the $5d$ coupling is expected to be quite small one would expect little variation with the group involving the common series limit. Clearly either the interval factor for $5d^3F_2$ is too small or that for $5d^1D_2$ is too large.

The conclusions can be summarized thus:

- (1) The coupling constant of the $7s$ electron is much smaller than that of $6s$.
- (2) In the (4S) system the $5p^3$ group can couple to form a constant comparable with that of the $6s$ electron.

(3) In the (2D) system the $5p^3$ group can probably couple to form a constant appreciably larger than in (4S).

(4) Three terms have interval factors larger than expected.

(5) In at least two terms the coupling constant is much smaller than it should be.

It is shown in the next section that the last two conclusions may be accounted for in terms of perturbations.

Part A.

Perturbations

The study of perturbations in the intervals of the fine structures of spectral terms has recently acquired importance owing to the discovery of Schüler and co-workers (Schüler and Schmidt 1935) who pointed out that a breakdown in the interval rule can occur in certain terms if the electric charge distribution is not spherically symmetrical, that is if the nucleus exhibits electrical quadrupole moment. The interaction between a non-spherically symmetrical nucleus with a non-spherically symmetrical electron charge distribution obeys the law $E = a \cos IJ + b \cos^2 IJ$ instead of the simple cosine law, as Schüler has pointed out. The constant b is a measure of the departure of the nucleus from spherical symmetry. A detailed wave-mechanical formula relating the observed constant b with the nuclear electrical distribution has been given by Casimir (Casimir 1935) and by its aid it is possible in favourable cases to calculate the nuclear quadrupole moment.

Great care must be used before a perturbation can be attributed with certainty to the nucleus, since four distinct kinds of perturbation can arise.

(1) Perturbation due to nuclear quadrupole moment (Schüler and Schmidt 1935).

(2) Configuration interaction perturbation arises in a resonance form between two terms with the same L , S , J , and parity; even if the terms are far apart (Shenstone-Russell perturbation) (Shenstone and Russell 1932).

(3) When the gross structure multiplet separation is comparable with the fine structure separation due to nuclear moment, the intervals in the latter are distorted (Paschen 1932).

(4) Term interaction perturbation arises when two terms fall accidentally close together (Schüler and Jones 1932).

We wish at first to draw particular attention to configuration interaction perturbation which can still be considerable even when the interacting terms are far apart (exceeding say $10,000 \text{ cm.}^{-1}$). Since in intermediate

coupling, L and S have no precise meaning, it is not possible to state a definite selection rule governing the perturbations for this type of coupling. As the I^+ spectrum is extremely rich in terms, not far removed from each other, it is quite certain that configuration interaction must occur frequently. In effect, the wave functions characteristic of the two involved terms are mixed up, the result being that the terms partake of the properties of one another. This is of fundamental importance to fine structure analysis, for if a term with an inherent wide structure, due say to a single penetrating electron in the configuration, perturbs a term with an inherent narrow structure, then the fine structure in the former is diminished, whilst the latter term will exhibit an anomalously large fine structure interval factor.

Thus the first effect of configuration interaction is to produce anomalies in the values of the observed fine structure interval factors. We are unable to say whether the interval rule will still hold accurately in such perturbed terms. This question is being examined theoretically by Professor Hartree. It is an important point. If there is a small second-order effect in the interaction causing a slight breakdown in the interval rule, then great caution must be exercised in deriving a quadrupole moment. In fact it will at present almost be impossible to separate the effects. Furthermore, since the perturbations will be in opposite senses in the two terms, one term would give a positive and the other a negative nuclear quadrupole moment, if this effect were not recognized. (Cases in the literature have been reported where one term gives an apparent positive quadrupole moment and another term an apparent negative value.)

In Table IV are shown only some of the extremely large number of terms which the Hund theory predicts for I^+ . In addition to those shown there are

TABLE IV. TERMS IN I^+

Electron con- figuration	Basic terms of the I ⁺⁺ ion											
	⁴ S		² D						² P			
5s ² 5p ⁴	³ P ¹ D ¹ S											
5s 5p ⁵	³ P ³ S ¹ P ¹ S											
5s ² 5p ³ ns	⁵ S ³ S	³ F	¹ D						³ P	¹ P		
5s ² 5p ³ np	⁵ P ³ P	³ F ³ D	³ P	¹ F	¹ D	¹ P				³ D ³ P ³ S	¹ D ¹ P ¹ S	
5s ² 5p ³ nd etc.	⁵ D ³ D	³ G ³ F	³ D ³ P	³ S	¹ F	¹ G	¹ D ¹ P ¹ S				³ F ³ D ³ P	¹ F ¹ D ¹ P

whole systems of terms based upon the $5s\ 5p^4\ ns$, $5s\ 5p^4\ np$, $5s\ 5p^4\ nd$, etc., electron configurations. It is quite obvious that configuration interaction must be very frequent indeed, since large numbers of these terms fall close together. This fact can easily account for the existence of the anomalously large and small observed fine structure interval factors. To take a particular case. If a term interacts with one containing an unbalanced $5s$ electron the structure in the latter will be so great that, partaking of the properties of both, it will acquire a large interval factor. Conversely the structure in the perturbing term will diminish. In a similar manner the anomalous small interval factors, as in $(^4S)\ 7s\ ^3S_1$, will result if a term is perturbed by one with an inherently very small structure.

Part B.

In two known terms, and in one unidentified term with $J = 1$, a breakdown in the interval rule has been established. This is shown in fig. 11, the thick lines being the observed positions of the fine-structure levels in the terms, those calculated from the Landé interval rule being dotted. It is probable that at least in $(^2D)\ 6s\ ^3D_2$ the breakdown in interval rule is due to nuclear quadrupole moment, although we cannot entirely exclude the possibility of perturbation by an accidentally close-placed term with $J = 1$. (This would displace only the $F = \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$ terms as is observed.)

Assuming for the moment that nuclear quadrupole moment is responsible for the three perturbations, the interaction energy is of the form

$$E = a_0 + \frac{a}{2} \cdot C + b \cdot C \cdot (C + 1).$$

The constant a represents the ordinary unperturbed interval factor and the constant b is a measure of the nuclear electrical quadrupole moment, q . In favourable cases q can be calculated from a formula given by Casimir which is

$$q = \frac{-b \cdot 8 \cdot I \cdot (2I - 1), J \cdot (2J - 1)}{3e^2 \frac{1}{r^3} (3 \cos^2 d - 1)}.$$

Unfortunately it does not yet appear possible for us to evaluate the term $\frac{1}{r^3} (3 \cos^2 d - 1)$ so that we cannot make an estimate of the value of the quadrupole moment. The difficulty arises from the fact that the theoretical coupling constants have not been determined.

We are able, however, to calculate the value of b for three terms, namely, $(^3D) 6s\ ^3D_2$, $(^4S) 5d\ ^3D_2$ and X_1 . The latter term is unknown but is derived from the line 6161.9. This is unclassified. It exhibits a sharp triplet structure on a widescale, and it is quite clear that the J value of the term with structure

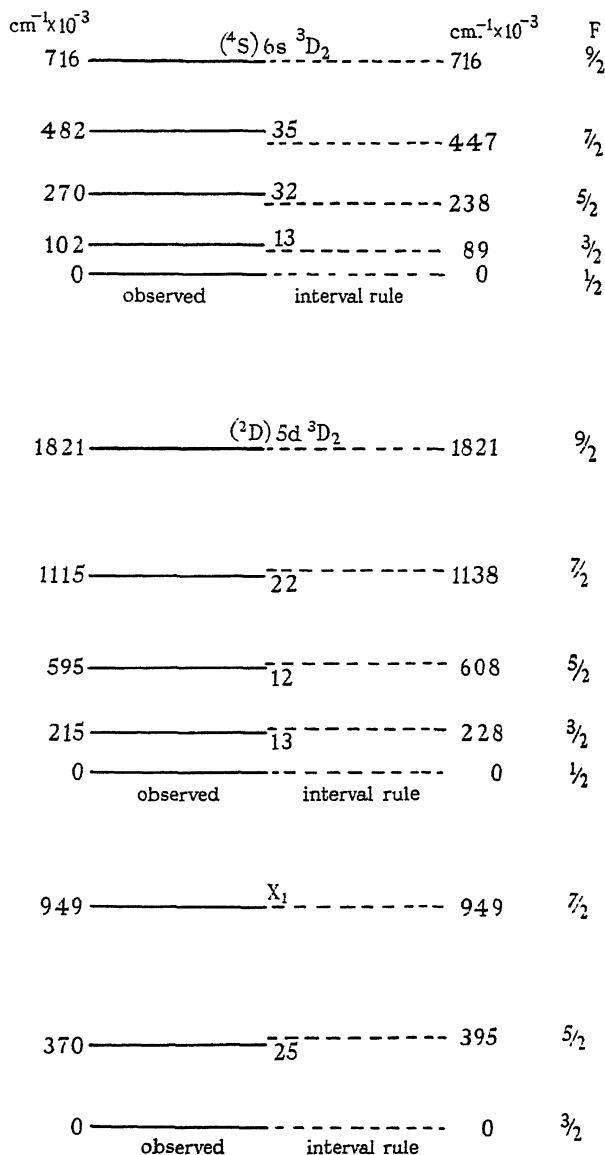


FIG. 11

must equal 1. The interval rule is widely departed from in this line and this cannot be attributed to the structure in the other term, for then the individual components would require to be very much broader than those observed.

The actual deviations from the interval rule are quite small. The best fit is given by the constants in Table V. In the terms with $J = 2$ the fit to the formula is only approximate.

TABLE V. CONSTANTS OF THE PERTURBED TERMS IN $\text{cm.}^{-1} \times 10^{-3}$

Term	a_0	a	b
$(^2\text{D})\ 6s\ ^3\text{D}_2$	454	58.4	-0.28
$(^4\text{S})\ 5d\ ^3\text{D}_2$	1039	152	+0.2
X_1	527.5	159	+0.73

The numerical ratio of a to b is exactly the same in the $(^2\text{D})\ 6s\ ^3\text{D}_2$ and X_1 terms.

It is of interest to note that no perturbation has been detected in the accurately measured $(^4\text{S})\ 6s\ ^5\text{S}_2$. This is in accordance with theory since the electrical distribution for this term is spherically symmetrical and it is only the interaction between an unsymmetrical nucleus with an unsymmetrical electron distribution which distorts the interval rule.

Corrections to multiplet classification

As previously stated a number of minor errors occur in the paper of Lacroute upon the multiplet classification of the lines. We have detected the following errors amongst the lines examined by us for fine structures:

The line 5405.3 given as $(^4\text{S})\ 5d\ ^3\text{D}_2-(^4\text{S})\ 6p\ ^3\text{P}_3$ should read instead $(^4\text{S})\ 5d\ ^3\text{D}_2-(^4\text{S})\ 6p\ ^5\text{P}_3$. The line 4488.5 given as $(^2\text{D})\ 5d\ ^3\text{D}_1-(^2\text{D})\ 6p\ ^3\text{D}_1$ should read instead $(^4\text{S})\ 5d\ ^5\text{D}_1-(^2\text{D})\ 6p\ ^3\text{D}_1$. Also the line 3302.4 given as $(^4\text{S})\ 5d\ ^3\text{D}_2-(^2\text{D})\ 6p\ ^3\text{P}_1$ should read $(^4\text{S})\ 5d\ ^3\text{D}_2-(^2\text{D})\ 6p\ ^3\text{D}_2$. Finally the value given for the term $(^4\text{S})\ 5s^2\ 5p^3\ 6p\ ^5\text{P}_1$ is 56885.21 cm.^{-1} . This should however read 56855.21 cm.^{-1} .

We wish to draw attention to the fact that the unclassified line 3833.7 has a fine structure remarkably similar to those involving the $(^4\text{S})\ 6s\ ^5\text{S}_2$ term. The structure is degraded to the red. It is therefore possible that this line results from the combination between this term and a deeply lying term with very small fine structure. If this is true, the position of the new term can be calculated. It is found to lie at 101,126 cm.^{-1} . The deepest term in the spectrum lies at 156,082 cm.^{-1} .

NOTE

In addition to the lines reported in Table I, eight others appeared upon a photograph taken with the 21 ft. grating. They were emitted only during the first few hours of life of the tube. Mr H. Bell has very kindly measured these plates for us. We give his results in Table VI.

The accuracy of measurement of the fine structures of these lines is very much less than that in Table I. Although a quintet pattern can be predicted for all the lines, this was only approximately resolved in two lines. We have reported the separations of as many components as are visible and then given the tail end of the pattern, so that in each case the total width is known approximately. These widths vary from 1.21 to 2 cm.^{-1} for the seven lines involving the $(^4\text{S}) 5d^3\text{D}_2$ term. Since the total width in the structure of the common term is 1.821 cm.^{-1} the observed structures are quite consistent with this when account is taken of the wide variation in the values of the upper interval factors.

Within the limits of measurement, all the lines have approximately the correct widths as calculated from the known interval factors of the terms involved. The only term amongst these lines, not common to lines in Table I is $(^2\text{D}) 6p^3\text{P}_1$ for which we get an interval factor of $\sim 37 \text{ cm.} \times 10^{-3}$ from the line 2925.

TABLE VI. APPROXIMATE FINE STRUCTURE MEASUREMENTS
MADE WITH A 21 FT. GRATING

Wave-length	Classification	Structure cm. $\times 10^{-3}$					Order
		Red		Violet			
5739.1	$(^4\text{S})\ 5d^3\text{D}_2 - (^4\text{S})\ 6p^5\text{P}_2$	0	619	1142.....	1650		I
		(5)	(4)	(3)	(1)		
5405.3	$(^4\text{S})\ 5d^3\text{D}_2 - (^4\text{S})\ 6p^5\text{P}_3$	0	631.....		1840		I
		(5)	(4)		(1)		
4828.3	$(^4\text{S})\ 5d^3\text{D}_2 - (^4\text{S})\ 6p^3\text{P}_2$	0.....			1760		I
		(5)			(1)		
4464.3	$(^2\text{D})\ 6s^3\text{D}_2 - (^2\text{D})\ 6p^3\text{P}_1$	0.....	650				II
		(5)	(1)				
3401.5	$(^4\text{S})\ 5d^3\text{D}_2 - (^2\text{D})\ 6p^3\text{F}_2$	0.....			1550		IV
		(5)			(1)		
2996.7	$(^4\text{S})\ 5d^3\text{D}_2 - (^2\text{D})\ 6p^3\text{F}_3$	0.....			< 2000		II
		(5)			(1)		
2957.5	$(^4\text{S})\ 5d^3\text{D}_2 - (^2\text{D})\ 6p^3\text{P}_2$	0	400	800	1110...1210		III
		(5)	(4)	(3)	(2)	(1)	
2925.1	$(^4\text{S})\ 5d^3\text{D}_2 - (^2\text{D})\ 6p^3\text{P}_1$	0	620	1146	1333...1600		III
		(5)	(4)	(3)	(2)	(1)	

This investigation has been rendered possible by assistance given by the Government Grant Committee to one of us (S. T.). The Lummer plate and other optical components were purchased with the grant received.

Our thanks are due to Mr H. Bell for permission to use his measurements.

SUMMARY

Earlier measurements by Tolansky upon the fine structures in the lines of the first-spark spectrum of iodine are extended into the ultra-violet region with the aid of a quartz Lummer plate interferometer. Fine structures are now known for fifty-eight lines, thirty-nine of which have been allocated to terms which go either to the (^4S) or the (^2D) series limit.

An independent fine-structure analysis is made for both the systems and is confirmed by the aid of intercombination lines in the ultra-violet. Fine-structure interval factors are given for 13 (^4S) terms and 15 (^2D) terms. These exhibit anomalies which can be explained as arising from configuration interaction perturbation.

It is shown that the nuclear coupling of the $5p^3$ electrons in iodine is comparable with that of the $6s$ electron.

In three terms, deviations from the interval rule have been established. These are attributed to nuclear electrical quadrupole moment. The perturbation constants have been calculated, but these cannot as yet be employed to give a numerical estimate of the quadrupole moment, since the coupling constants have not yet been evaluated.

[*Note added in proof*, 24 August 1938. After communicating these results, Murakawa in *Z. Phys.* 109, 162, 1938, published an extensive gross structure multiplet classification differing in some respects from that of Lacroute. Almost all the strong unclassified lines in Table I are allocated and a few term identifications are altered, the J values being however retained. Lacroute's (^4S) $5d\ ^3\text{D}_2$ is now altered to $5s\ 5p^5\ ^3\text{P}_2$, removing thus the anomaly regarding the large interval factor. Eight terms are perturbed, those with $J = 2$ failing to fit Schüller's perturbation formula, obeying instead $E = a_0 + \frac{a}{2} \cdot C + b \cdot C(C+1) + C \cdot C^2(C+1)$. The last term is attributed to a new phenomenon, namely nuclear *magnetic* quadrupole moment. A detailed analysis is being communicated as Part IV.]

REFERENCES

- Casimir 1935 *Physica*, 7, 719.
 Fermi and Segré 1933 *Z. Phys.* 82, 729.
 Lacroute 1934 Theses. University of Paris, Nov.
 Paschen 1932 *S.B. preuss. Akad. Wiss.* 32, 505.

- Schüler and Jones 1932 *Z. Phys.* **77**, 801.
Schüler and Schmidt 1935 *Z. Phys.* **98**, 239.
Shenstone and Russell 1932 *Phys. Rev.* **39**, 415.
Tolansky 1935*a* *Proc. Roy. Soc. A*, **149**, 269.
— 1935*b* *Proc. Phys. Soc.* **48**, 49.
White 1930 *Proc. nat. Acad. Sci., Wash.*, **16**, 68.
-

A new process of negative-ion formation. IV

BY F. L. ARNOT, PH.D.

Lecturer in Natural Philosophy

AND CLARK BECKETT, B.SC.

The University, St Andrews

(Communicated by H. S. Allen, F.R.S.—Received 22 June 1938)

INTRODUCTION

The three previous papers of this series (Arnot and Milligan 1936*b*; Arnot 1937*a, b*) contain an account of experimental work which led the senior author to propose a new process of negative-ion formation. This process is the formation of negative ions at metal surfaces by bombardment of the surface with positive ions, the negative ion being formed by the positive ion capturing two electrons from the surface.

Further work carried out during the past year, which is described in this paper, has revealed a new variation of the above process. In this latter process the impinging positive ion causes an adsorbed atom on the surface to come off as a negative ion. It is believed that this newer process is essentially similar to the process previously reported, the difference being due merely to the transference of excitation energy from the incident positive ion, after its capture of an electron, to the atom adsorbed on the surface. The discovery of this second new effect was made independently by Sloane and Press (1938), although they attribute it to a different process.

APPARATUS AND EXPERIMENTAL PROCEDURE

The apparatus used in this work, which is shown in fig. 1, has been fully described in three previous papers (Arnot and Milligan 1937*a, b*; Arnot 1937*a*). It will be sufficient to repeat here that it consists of an ionization chamber and a magnetic analyser. The analyser, which is evacuated through a liquid-air trap of large diameter by the right-hand outlet, was set in a gap of 1 cm. between the poles of a large electromagnet. The left-hand outlet was connected through another liquid-air trap to a McLeod gauge.

All the results given in this paper were obtained with a quartz ionization chamber No. 2 shown as an inset in fig. 1. This contained two nickel gauze electrodes and a tungsten filament of about 1 cm. in length which could be withdrawn for renewal by means of a ground joint. The upper gauze electrode will be denoted by E_1 and the lower one by E_2 . The difference of potential between the filament and E_1 will be denoted by V_0 , that between E_1 and E_2 by V_1 , and that between E_2 and the iron cylinder T by V_2 . For all the results given in this paper the negative ions entered the analyser with an energy of 200 eV. The negative ions produced on the filament were obtained by setting $V_0 = 200$ V, V_1 and V_2 both being zero. The negative ions produced on E_2 were obtained by setting $V_0 = 100$ V, $V_1 = -200$ V, $V_2 = 200$ V. The negative sign attached to V_1 means that this potential was in a direction to *retard* electrons or negative ions coming from the filament. The soft iron shielding cylinder and plate shown in fig. 1 were not used in these experiments, and consequently the electrons from the filament were prevented by the stray field of the magnet from penetrating more than a short distance through E_1 . The positive ions produced between E_1 and E_2 will thus all arise close to E_1 and will pass through E_2 with an energy of nearly 200 eV. On entering the field V_2 these positive ions will be stopped and reversed back through E_2 , where they will again be reversed by the field V_1 . They will consequently pass back and forth through the gauze E_2 , some of them being absorbed by the gauze in each passage of the ions through it. Thus E_2 will be bombarded by positive ions equally on both sides. The negative ions produced on the under side of E_2 will then enter the analyser with an energy of 200 eV. With this arrangement of potentials the negative ions produced on the filament will all be stopped between E_1 and E_2 . No negative ions are emitted from E_1 , since this gauze is not subjected to positive ion bombardment. The production of negative ions from excited and metastable atoms is known to be ineffective (Smith 1938), and any produced on E_1 will be driven into the electrode by the fields V_0 and V_1 .

The gas to be investigated was stored in a large glass bulb of about $1\frac{1}{2}$ l.

capacity at a pressure of a few centimetres of mercury. The gas from the bulb passed through a fine glass capillary, then through a liquid-air trap, and entered the ionization chamber through a side tube sealed on in place of that containing mercury in fig. 1. After passing through the apparatus the gas was pumped off into the atmosphere. The pressure of gas in the ionization chamber was usually adjusted to a value between 2×10^{-4} and 5×10^{-4}

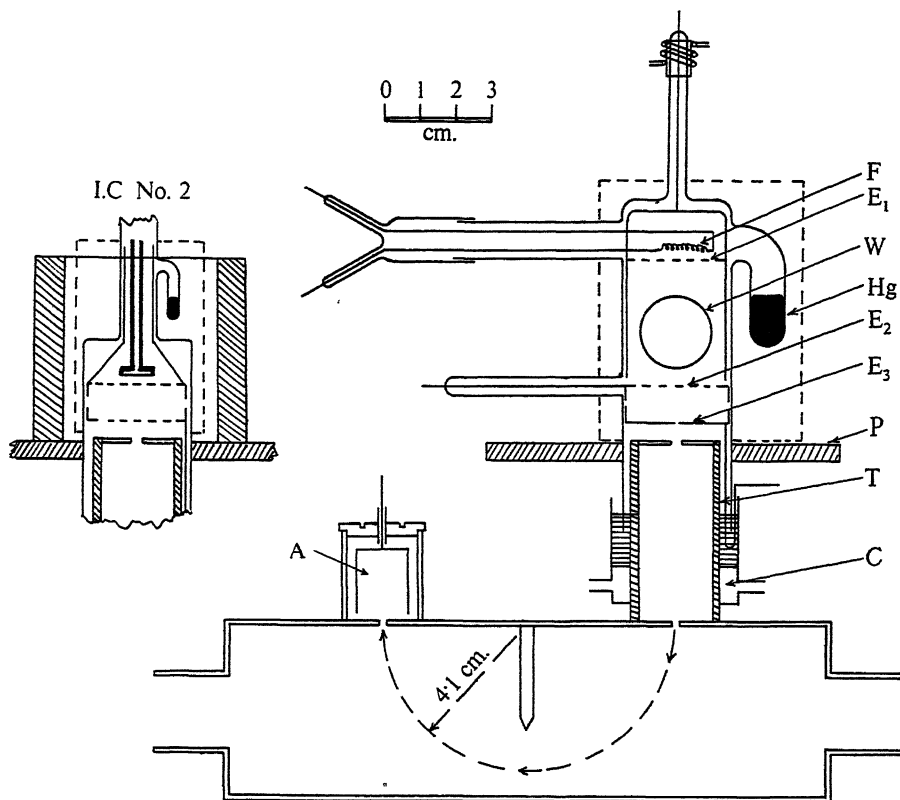


FIG. 1. Apparatus.

mm. of Hg, and this value remained constant during a run. At one period of the work a second McLeod gauge was connected directly to the ionization chamber. Comparison of the reading of this gauge with the reading of that connected to the analyser showed that the pressure in the analyser was only about 5 % less than the pressure in the ionization chamber. The argon used was purchased in a glass cylinder from the British Oxygen Co., and was stated by them to be spectroscopically pure. The oxygen and nitrogen were both obtained from commercial cylinders newly filled. All taps and ground

joints in the vacuum system were greased with Apiezon grease of negligible vapour pressure.

All the results given in this paper were obtained with a filament emission of 3 mA, except the results shown in the lower curve of fig. 4 for which the emission was 10 mA. When the potentials were set to produce negative ions on E_2 the positive ion current to E_2 was of the order of 10^{-5} amp. except, of course, when runs were taken in vacuum.

For all vacuum runs the positive ion current to E_2 was undetectable on the microammeter which could easily detect 10^{-7} amp. Since E_2 acts as the collector of an ionization pressure gauge we may conclude that the pressure of gases and vapours in the ionization chamber during a vacuum run was less than a hundredth of the pressure prevailing for a run in gas, that is, less than 5×10^{-6} mm. of Hg. The pressure on the McLeod gauge was always undetectable during a vacuum run.

All results for negative ions from E_2 obtained in gas, except those of fig. 3, have been reduced to a constant positive ion current to E_2 of 2×10^{-6} amp. The positive ion current to E_2 was unfortunately not measured for the earlier results given in fig. 3. The ordinate scale of all curves is marked in units of 10^{-14} amp. When a peak is too strong to be plotted completely on the scale adopted for a figure the ordinates have been reduced by a factor of ten, and the peak plotted as a broken curve.

The magnet was calibrated for both positive and negative ions by means of a search coil and Grassot fluxmeter, which was checked for linearity of scale by a ballistic galvanometer. The fluxmeter was also checked against a Hibbert standard flux. The calibration curve was found to be linear except for a short range near the origin. From this curve a curve was obtained of mass number as a function of magnet current. All the figures in this paper are plotted with an abscissa scale of mass number, the O^- peak being set at mass number 16. The mass numbers of the peaks obtained by varying the magnet current were checked in a typical curve by keeping the magnet current constant and varying the energy of the ions entering the analyser. The mass numbers of the peaks obtained by both methods were in good agreement.

More than seventy runs have been taken. Results obtained under the same experimental conditions are completely consistent even in the finer details, and each curve given in this paper has been checked several times.

RESULTS

(1) *Positive-ion analysis*

A typical positive-ion analysis of the ions formed in argon by electrons of 100 eV energy is shown in fig. 2. This is in good agreement with the

results of Bleakney (1930), who finds the ratio of the number of Ar^{++} ions to the number of Ar^+ ions to be 0.103 for an electron energy of 100 eV. This ratio obtained from the peaks in fig. 2 is 0.108. In calculating this ratio the size of the Ar^{++} peak must, of course, be halved since each Ar^{++} ion carries two charges. In addition to the two argon peaks at mass numbers 20 and 40 there are traces of N^+ at 14, O^+ and H_2O^+ at 16–18, and N_2^+ or CO^+ at 28. The lower curve in fig. 2 is a vacuum run plotted to the same scale. It is evident from this that the above impurities are brought out of the glass and electrodes by the intense positive ion bombardment.

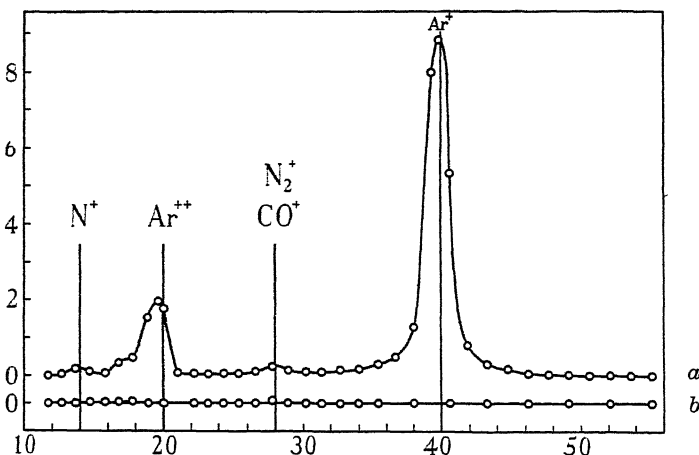


FIG. 2. (a) Argon positive-ion analysis. (b) Vacuum run.

(2) Negative-ion analysis

Three typical analysis curves for negative ions produced on the gauze E_2 are shown in fig. 3. For the upper curve E_2 was bombarded with positive ions of argon; for the second curve nitrogen positive ions were used; and for the lower curve E_2 was bombarded with positive ions of mercury by removing the liquid-air traps.

In their main features these three curves are all very similar. We see that a large peak of negative ions of mass number 24 is produced by each type of positive ion. In addition to this strong peak there is a relatively faint background of negative ions of mass numbers 12, 16, 29, 32, 38, 44 and 51. The spectra of negative ions obtained from the filament were essentially of the same type as those obtained from E_2 , the peak of mass number 24 again being by far the strongest peak.

When E_2 or the filament is bombarded with positive ions of oxygen entirely different negative ion spectra are obtained as is shown by fig. 4, in which the upper curve is for ions from the filament and the lower curve is for

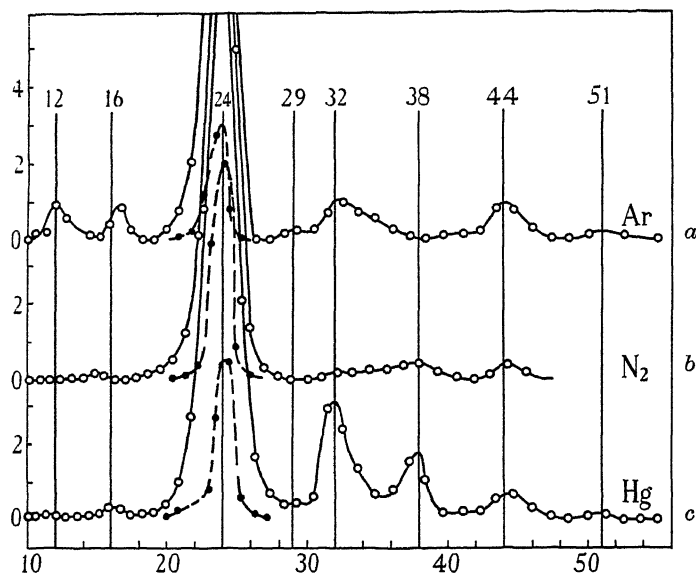


FIG. 3. Analysis of negative ions formed at E_2 by positive ions of (a) argon, (b) nitrogen and (c) mercury.

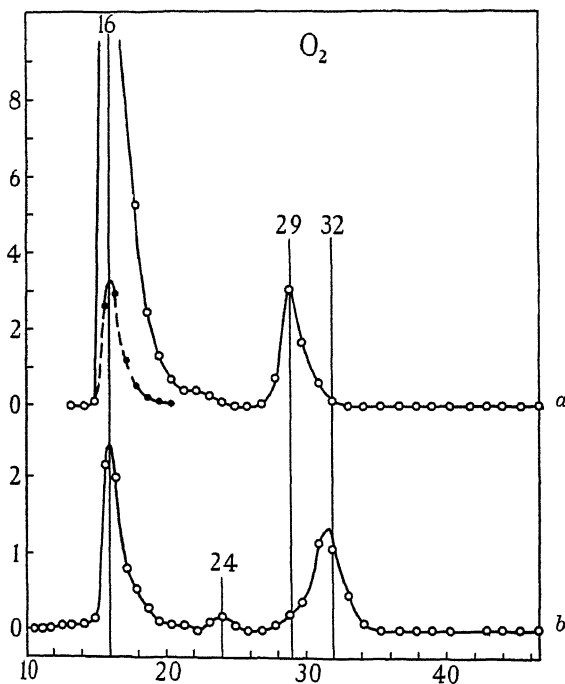


FIG. 4. Analysis of negative ions formed by positive ions of oxygen (a) at the filament, (b) at E_2 .

The oxygen was now all pumped out of the apparatus and a vacuum run taken. We see from this curve, which is plotted at the top of fig. 5 as curve (a), that no appreciable peaks were obtained. By removing the liquid-air traps so as to bombard E_2 with positive ions of mercury we might expect to

reproduce curve (c) of fig. 3 which represents the spectrum of negative ions obtained from E_2 by mercury positive ions before the runs in oxygen had been taken. However, the curve actually obtained, which is shown in fig. 5 as curve (b), is quite different from curve (c) of fig. 3. The only difference in the experimental conditions under which the results shown in figs. 3 (c) and 5 (b) were obtained is that, in between these runs, oxygen had been admitted

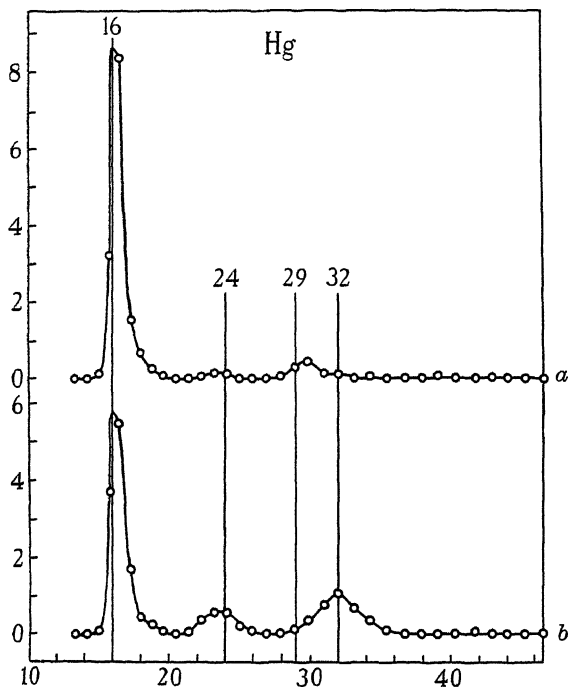


FIG. 6. Analysis of negative ions formed at the filament by positive ions of mercury. Curve (a) was obtained with an oxidized filament. After cleaning the filament by flashing it, curve (b) was obtained.

to the apparatus. The effect of this is to suppress the peak at 24. This is probably caused by the retention of adsorbed oxygen on the surface of E_2 , for when the surface of E_2 is fully covered with oxygen we have seen from fig. 4 (b) that the peak at 24 is very small indeed. We shall see later from fig. 8 that when the adsorbed oxygen has been removed from E_2 by persistent positive-ion bombardment the peak at 24 returns to its full strength. A curve practically identical to that of fig. 5 (b) was obtained by using, instead of Hg^+ ions, argon positive ions with liquid air on all three traps.

The spectrum of negative ions obtained from the filament by bombardment with positive ions of mercury after the oxygen had all been pumped out is shown in fig. 6 (a), which is similar to the curve taken in oxygen (fig. 4 (a)).

The hot filament will have absorbed a considerable quantity of oxygen while running in this gas, and consequently its surface will be well covered with oxygen which is being desorbed from the filament when running in mercury vapour. The similarity of figs. 6 (*a*) and 4 (*a*) shows that O^- ions are produced by bombardment of the filament by positive ions of mercury as well as by positive ions of oxygen. We notice again from figs. 6 (*a*) and 4 (*a*) that the peak of mass number 24 is very small when the surface is well covered with oxygen. Curves practically identical to that of fig. 6 (*a*) were obtained by using, instead of Hg^+ ions, argon and nitrogen positive ions with liquid air on all three traps.

The filament was then flashed, that is raised to a high temperature for a few seconds, and the runs of figs. 5 (*b*) and 6 (*a*) repeated. These are shown in figs. 5 (*c*) and 6 (*b*). Flashing of the filament reduces its content of absorbed oxygen and so reduces the O^- as shown in fig. 6. The oxygen given off when the filament is flashed increases the amount of oxygen adsorbed on E_2 resulting in a decrease of the peak at 24 and an increase of the O_2^- peak, as shown in fig. 5. The mercury was then frozen out with liquid air and another vacuum run taken. This is shown as curve (*d*) in fig. 5.

We have seen that the peak of mass number 24 is suppressed by the absorption of oxygen by the surface on which the negative ion is formed. By persistent bombardment of E_2 with positive ions we should get rid of a good deal of this adsorbed and occluded oxygen, and we should then expect to obtain curves similar to those shown in fig. 3 which were obtained before oxygen had been admitted to the apparatus. The set of curves shown in fig. 7 shows that this is realized. Curve (*a*) is a vacuum run. Curve (*b*) is a run in nitrogen, and curve (*c*) is another vacuum run taken after the nitrogen had been pumped out. Curves (*d*), (*e*) and (*f*) represent corresponding runs in argon. The nitrogen curves of figs. 3 (*b*) and 7 (*b*) are similar in that the main peak is now the ion of mass number 24. The O_2^- peak is, however, more prominent in the later curve owing to the surface not being so free of oxygen as it was when the earlier curves were taken. The argon curves of figs. 3 (*a*) and 7 (*e*) are also similar, except that here again the O_2^- peak is more prominent in the later curve than in the earlier curve. Argon was always found to bring out the O_2^- peak from E_2 rather more strongly than did nitrogen, as is shown by figs. 3 and 7. It will also be observed from curves (*a*), (*b*) and (*c*) of fig. 7, which were taken in this order, that bombardment of E_2 with positive ions reduces the amount of adsorbed oxygen on the surface and so reduces the O^- and O_2^- peaks.

The analysis of negative ions from the filament produced by bombardment with positive ions of nitrogen and argon is shown in fig. 8. Curves (*a*) and (*c*)

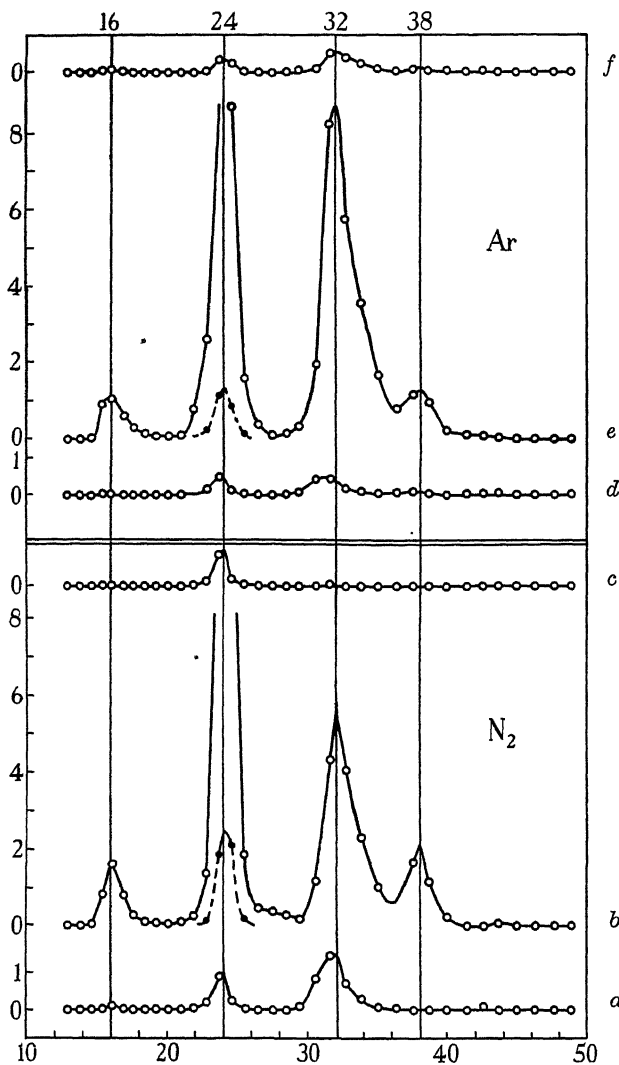


FIG. 7. Analysis of negative ions formed at E_2 by positive ions of (b) nitrogen, (e) argon. Curves (a) and (c) are vacuum runs taken respectively before and after the run in nitrogen. Curves (d) and (f) are vacuum runs taken respectively before and after the run in argon.

are vacuum runs taken before and after the run (b) in nitrogen. Curves (d) and (f) are vacuum runs taken before and after the run (e) in argon. Curve (g) is a repetition of part of the vacuum run (f) taken after an interval of an hour. We see from these curves in nitrogen and argon that the O^- peak is still quite strong, showing that absorbed oxygen is still diffusing to part of

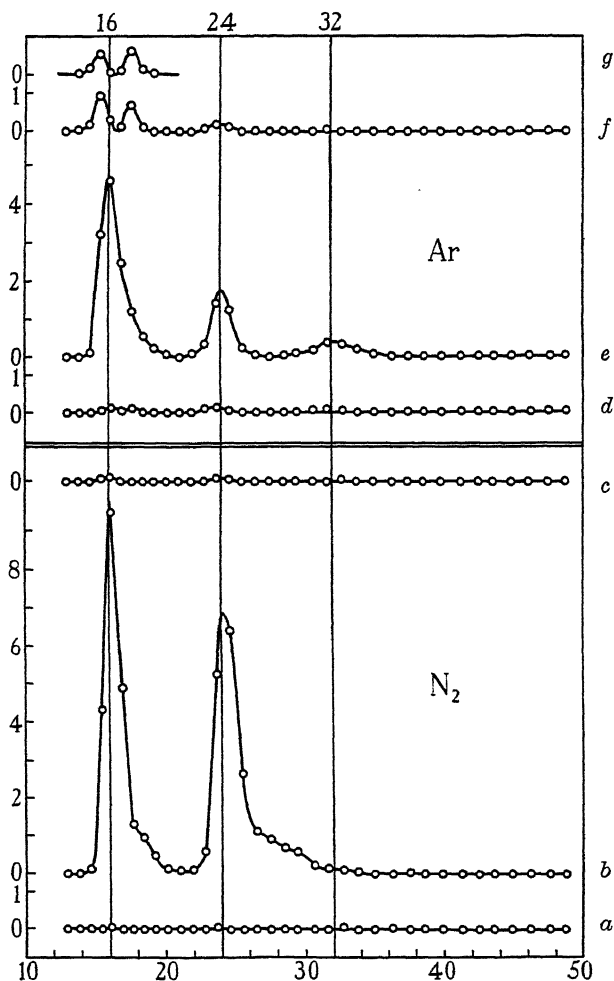


FIG. 8. Analysis of negative ions formed at the filament by positive ions of (b) nitrogen, (e) argon. Curves (a) and (c) are vacuum runs taken respectively before and after the run in nitrogen. Curves (d) and (f) are vacuum runs taken respectively before and after the run in argon. Curve (g) is a repetition of part of curve (f) taken an hour later.

the surface of the filament. However the peak of mass number 24 is considerably stronger in these curves than in fig. 6 owing to a larger area of the filament surface being free of oxygen in these later curves. Owing to the large temperature gradient from the centre to the ends of the short 1 cm. filament it is probably impossible to free the whole filament from absorbed oxygen by flashing.

DISCUSSION

In the first experiments on this work (Arnot and Milligan 1936*b*; Arnot 1937*a, b*) it was found that when a surface was bombarded with positive ions of one species negative ions of the same species were emitted from the surface. For example, when the surface was bombarded with Hg^+ ions Hg^- ions were emitted. To account for this new process of negative ion formation Arnot proposed the following mechanism. The positive ion first captures an electron from the surface into an excited state. This excited atom then captures a second electron from the surface thus becoming a negative ion. The double capture process will be energetically possible provided

$$V_i + V_f > 2\phi, \quad (1)$$

where V_i is the ionization potential and V_f the electron affinity of the incident atom, and ϕ the work function of the surface. Our colleague, Dr R. A. Smith (1938), has worked out the details and the quantum mechanics of this process, and he finds that the theory accounts very satisfactorily for the experimental results. An important result found by Smith is that there exists a high probability of the newly formed negative ion being neutralized by the electron being captured back into a vacant state in the metal by a resonance process. This explains why the probability of negative-ion formation by this process increases, instead of decreases, with increase in the velocity of the incident positive ions, since the probability of neutralization of the newly formed negative ion is less the shorter the time spent in the vicinity of the surface. The actual probability of formation of the negative ion from the positive ion is very high, but the large majority of these newly formed negative ions are neutralized before they escape from the neighbourhood of the surface.

Since the negative ion is formed in two stages the energy condition (1) can be separated into two, the first applying to the capture process which neutralizes the positive ion and the second to the capture process which converts the neutral excited atom into a negative ion. Both these capture processes are probably due to resonance between the energy states of the atom and those of the surface. We thus have

$$V_i - V_e > \phi, \quad (2)$$

$$V_e + V_f > \phi, \quad (3)$$

where V_e is the energy of the excited state into which the first electron is captured.

The results presented in this paper show that positive ions of one element may give rise on impact with a surface to negative ions of a different element. In a preliminary report of this work (Arnot and Beckett 1938) we suggested that this new process may be due to a collision of the second kind between the incident positive ion and an atom of gas or vapour adsorbed on the surface. Thus we suppose the incident positive ion to capture an electron into an excited state by resonance with the surface, the condition for this being given by (2). This excited atom then makes a collision of the second kind with an atom on the surface, exciting it to a state V'_e which will normally be the state nearest in energy to the state V_e of the incident atom, the energy difference $V_e - V'_e$ going into kinetic energy if $V_e > V'_e$ or being made up from the kinetic energy of the incident atom if $V_e < V'_e$. The excited atom from the surface, on reverting to its normal state, captures an electron from the surface by a resonance process thus becoming a negative ion. This process is the same as that which accounts for those negative ions which are formed from their own positive ions, except that in the new process transfer of excitation and kinetic energy occurs between the two species of atoms. Transfer of kinetic energy from the incident atom to the adsorbed atom must take place, since the negative ion formed by this process has an energy in excess of that acquired in the accelerating field. We thus have

$$V_i - V_e > \phi, \quad (4)$$

$$V'_e + V_f > \phi, \quad (5)$$

$$V_e - V'_e \leftrightarrow \text{kinetic energy.} \quad (6)$$

V_i is the ionization potential of the incident atom as before, and V_f is the electron affinity of the atom adsorbed on the surface which comes off as the negative ion. V'_e should not in general differ by more than a fraction of a volt from V_e since the probability of a collision of the second kind taking place is only high for close resonance.

Returning now to the experimental results, we have seen that O_2^- always appears in the spectra of ions arising on E_2 when this electrode is bombarded by positive ions of oxygen, mercury, nitrogen or argon. This ion is not, however, formed by bombardment of an *oxidized* tungsten filament by positive ions of oxygen or mercury* as shown by figs. 4 (a) and 6 (a). We have also found that it is not formed by positive ions of argon or nitrogen so long as the filament is fully oxidized. We see from fig. 6 (b) that if the filament is cleaned by flashing the O_2^- peak appears.

* A very slight trace of O_2^- is observed in fig. 6 (a), but this is undoubtedly due to the filament not being fully oxidized over its entire length.

The explanation of this result may lie in the difference between the work functions of clean and oxidized tungsten. The values of the work function for clean and oxidized tungsten are given by Reimann (1935) as 4.64 and 6.34 V respectively. Massey (1937) quotes values for the electron affinity of O and of O_2 derived from the experiments of Lozier (1932), Loeb (1935) and Bradbury (1933). The value of V_f for O is 2.2 V and for O_2 the value lies between 0.1 and 0.3 V, the mean value being 0.2 V. From the condition (5) governing the capture of a second electron to form a negative ion we have

$$V'_e > \phi - V_f.$$

Thus for the formation of O_2^- the excitation energy of the O_2 molecule must be greater than $6.34 - 0.2 = 6.1$ eV if the formation takes place at an oxidized tungsten surface, and greater than $4.64 - 0.2 = 4.4$ eV if the capture occurs from a clean tungsten surface. It may be significant that the work of dissociation of O_2 , which is 5.1 eV, lies between these two values, and consequently O_2 may dissociate with or without capture at the oxidized surface, while at a clean surface it would form O_2^- .

Whether or not this is the correct explanation, there is no doubt about the experimental result that, whereas a strong peak of negative ions can be obtained at mass number 32 from a clean tungsten surface, this peak either completely disappears or is barely perceptible when the tungsten surface is oxidized, and is replaced by a fairly strong peak at 29. In view of the undoubted importance of this result from the theoretical viewpoint another set of curves demonstrating the effect is shown in fig. 9. Curve (a) which was taken in oxygen at a pressure of 4×10^{-4} mm. shows the O^- peak at 16 and the peak at 29. There is no sign of a peak at 32. The oxygen was then all pumped out and the filament flashed in vacuum to clean it. The vacuum run shown as curve (b) was then taken. The liquid-air traps were then removed to provide mercury positive ions for bombarding the filament and curve (c) was taken. In addition to the O^- peak there now appears a definite peak exactly at mass number 32 which must be O_2^- . The strong peak which formerly appeared exactly at mass number 29 has decreased in intensity by a factor of ten. Nine determinations of the mass number of this peak taken from nine separate curves are given in Table I together with the gas which provided the positive ions for bombardment of the filament.

All of these values are taken from curves obtained with an oxidized filament. In six of these curves the O_2^- peak was entirely absent; in the remaining three curves there was a very slight inflexion on the high mass side of the 29 peak at approximately 32 which could be accounted for by the filament not being fully oxidized.

We see from this table that there is no doubt that the peak appears at a mass number of 29 and cannot therefore be ascribed to CO^- . Moreover, in none of our curves have we found a CO^- peak, which is not surprising in view of the fact that N_2^- , which is believed to be unstable, has a similar outer electronic structure to CO^- . The nature of this peak at mass number 29 is very uncertain; all we can say at present is that its mass number corresponds to COH^- and to C_2H_5^- .

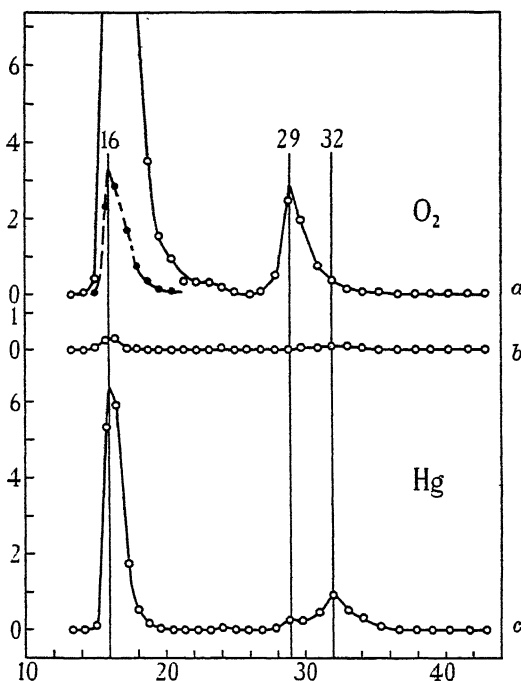


FIG. 9. (a) Analysis of negative ions formed at an oxidized filament by positive ions of oxygen. (b) Vacuum run taken after the filament had been cleaned by flashing in vacuum. (c) Analysis of negative ions formed at the clean filament by positive ions of mercury.

TABLE I

Oxygen	29.0	Nitrogen	28.3
Oxygen	28.8	Argon	29.6
Oxygen	29.2	Mercury vapour	29.3
Oxygen	28.9	Residual gas from an	
Oxygen	29.0	un-degassed filament	29.0

We see from figs. 4, 6, 8 and 9, in all of which the O^- peak is strong, that there appears a bulge or unresolved peak on the high energy side of the O^- peak. This was previously observed by Arnot (1937*a, b*) and was shown to be

due to fast O^- ions, which come off the surface with an energy of 35–70 eV when the parent positive ions had an energy of 200 eV. This excess energy means that the accommodation coefficient has a second principal value of about 0.715, the main principal value being 0.985. Arnot showed that this effect was most clearly demonstrated for H^- ions for which two well-resolved peaks were obtained.

The excess energy of the fast O^- ions calculated from the position of the bulge on the side of the main O^- peak in figs. 4 and 9 is about 70 eV. From fig. 6 the value is about 38 eV, and from the two well-resolved peaks of figs. 8 (*f*) and 8 (*g*) the value obtained is 34 eV.

The mass numbers of the peaks observed in this work are given in Table II together with the figure in which the ions appear. When the intensity of the peak is so low that it is just perceptible its mass number is enclosed in brackets. A dash in the columns of mass number 12 and 51 signifies that measurements were not extended down to, or up to, these mass numbers. With the exceptions of O^- and O_2^- the negative ions indicated at the foot of each column must be taken merely as tentative suggestions.

TABLE II

Fig.	Mass number of negative ions							
	12	16	24	29	32	38	44	51
3	12	16	24	29	32	38	44	51
4	12	16	24	29	32	—
5	—	16	24	...	32	...	(44)	—
6	—	16	24	29	32	(38)	...	—
7	—	16	24	...	32	38	(44)	—
8	—	16	24	29	32	(38)	...	—
9	—	16	(24)	29	32	—
Ion	C^-	O^-	C_2^-	COH^- $C_2H_5^-$	O_2^-	$C_3H_2^-$	CO_2^-	$C_4H_3^-$

The impurities detected in positive ion analysis curves were slight traces of N^+ , O^+ , H_2O^+ , CO^+ or N_2^+ and CO_2^+ having mass numbers 14, 16, 18, 28 and 44. With the filament properly run in, the electrodes bombarded, and liquid air on all three traps these peaks were imperceptible as shown by fig. 2 (*b*). When the electrodes were under intense positive-ion bombardment slight traces of some or all of the above impurities were brought out of the electrodes and appeared as shown by fig. 2 (*a*). No impurities of mass numbers 24, 29, 38 or 51 were ever detected by us in positive-ion analyses. The negative ions of these mass numbers could not therefore have been produced by bombardment of the electrodes by the corresponding positive ions in the same way as O^- and O_2^- are formed by bombardment with

positive ions of oxygen. These negative ions must arise from atoms and molecules adsorbed on the surface of E_2 and the filament. Similarly the O^- and O_2^- produced by bombardment with positive ions of mercury, nitrogen and argon must arise from adsorbed atoms and molecules of oxygen. We therefore conclude that there are two processes leading to the formation of negative ions at surfaces. The initial stage in each process is the same; this is capture of an electron from the surface by the incident positive ion into an excited state of the atom, probably by a resonance process. This excited atom, on reverting to its normal state, may then capture a second electron from the surface, again probably by a resonance process, so forming a negative ion of the same element as the incident positive ion. The second process consists in the excited atom transferring its excitation energy and part of its kinetic energy to an adsorbed atom on the surface which then captures an electron from the surface, and so forms a negative ion of an element different to that of the incident positive ion. The energy conditions for these two processes of formation have already been given on p. 115.

To prove that there were no condensable vapours in the ionization chamber or that, if there were, they had no effect on the experimental results we surrounded the ionization chamber from the first slit of the analyser up to the filament with a mixture of solid CO_2 and alcohol contained in a wooden trough built round the ionization chamber. On then taking a run in argon under the experimental conditions prevailing for curve (c) in fig. 5 a practically identical curve was obtained.

Smithells and Ransley (1936) have shown conclusively that the CO desorbed from nickel at high temperatures is derived from dissolved oxygen and carbon, and is not present as CO molecules in the metal. The dissolved oxygen and carbon present on the surface of the nickel electrode E_2 might thus account for the peaks at mass numbers 16 and 32, 12 and 24 obtained by bombardment with positive ions of mercury, nitrogen and argon. It has been observed by Mann (1937) that the surface of a platinum wire was far more resistant to contamination after it had been heated in oxygen. If this applies also to tungsten it would account for the suppression of the peaks of mass numbers 24 and 38 observed when the filament had been run in oxygen.

We must now review the previous work (Arnot and Milligan 1936*b*; Arnot 1937*a, b*) on the surface formation of negative ions in the light of these new results. The work on the formation of negative ions of mercury, hydrogen and oxygen requires no amendment except that the probabilities of formation of these negative ions per positive ion may be higher than previously reported owing to some of the positive ions being employed in producing

negative ions from adsorbed atoms on the surface. The present work shows that no stable negative ions of nitrogen or argon were observed, the N^- peak formerly reported being O^- produced by the positive ions of nitrogen by the new process shown to occur in this paper. It was shown in the previous work that the intensity of this peak was proportional to the total positive ionization, and not to the distribution of this positive ionization amongst the monatomic and diatomic forms. Consequently N^+ and N_2^+ ions must be equally effective in producing O^- ions by the new process.

The new calibration of the magnet shows that the peak marked CO^- in the previous work is really the ion of mass number 24 appearing in this work, and the peak marked CO_2^- is the peak of mass number 38 which also appears in this work.

The mass numbers assigned to the peaks in this work are based upon the identification of the strong peak in oxygen, fig. 4, as O^- . The calibration curve for the magnet then gives the mass numbers of the other peaks. There seems little doubt that the identification of the strong peak obtained in oxygen as O^- is correct, since the curves given in fig. 3 show that this peak was very small before oxygen was admitted to the apparatus. The only doubt we feel about the correctness of this identification is caused by the report of Sloane and Press (1938) that they obtain a strong peak of CO^- ions, whereas with this identification we obtain no trace of CO^- ions. If we accept the result of Sloane and Press, and identify the strong peak in fig. 3 as CO^- of mass number 28 instead of our value of 24, then the strong peak obtained by us in oxygen would have mass number 18.7 which does not correspond exactly with any known ion. The ion of nearest mass number is H_2O^- . This could not be formed directly from H_2O^+ ions because no change in the intensity of this peak occurred when the ionization chamber was cooled with a mixture of solid CO_2 and alcohol, as mentioned above. Neither could the H_2O^- ions be produced from adsorbed molecules of H_2O on the surfaces for we should then expect this peak to be just as strong in fig. 3 when positive ions of mercury, nitrogen and argon were employed for bombardment. Direct calculation of the $H\rho$ of the strong peak in fig. 3 gives the mass number as 25 which is nearer to 24 than to 28. Although there is little doubt that the identification of the strong peak obtained in oxygen as O^- is correct an independent determination with a mass spectrograph of higher resolving power would be very valuable.

One of us (C.B.) wishes to thank the Carnegie Trust for the Universities of Scotland for a Research Scholarship.

SUMMARY

The three previous papers of this series described a new process of negative-ion formation. This process is the formation of negative ions at metal surfaces by bombardment of the surface with positive ions, the negative ion being formed by the positive ion capturing two electrons from the surface. The present paper shows that a new modification of this process can also occur.

The experimental results show that a spectrum of negative ions is obtained when a surface is bombarded with positive ions of argon or nitrogen or mercury, and that practically identical negative ion curves over the range of mass numbers from 10 to 54 are obtained by each type of positive ion. The negative ions observed have mass numbers 12, 16, 24, 29, 32, 38, 44 and 51. It is concluded that these negative ions arise from atoms and molecules of gas and vapour adsorbed on the surface.

It thus appears that there are two processes leading to the formation of negative ions at surfaces. The initial stage in each process is the same; this is capture of an electron from the surface by the incident positive ion into an excited state of the atom, probably by a resonance process. This excited atom, on reverting to its normal state, may then capture a second electron from the surface, again probably by a resonance process, so forming a negative ion of the same element as the incident positive ion. The second process consists in the excited atom transferring by a collision of the second kind its excitation energy to an adsorbed atom on the surface, which then captures an electron from the surface, and so a negative ion of an element different from that of the incident positive ion is produced.

The previous results are reviewed in the light of this new work which shows that no negative ions of nitrogen are formed. Certain alterations in the interpretation of the CO_2 results must also be made. The previous work on the formation of negative ions of mercury, hydrogen and oxygen requires no amendment, except that the probabilities of formation of these ions may have to be raised.

It is shown that no O_2^- ions are formed at the surface of an oxidized tungsten filament, but that they readily appear when the filament is cleaned by flashing. It is suggested that this effect is due to the increased work function of the filament when oxidized. It is also found that a covering of oxygen on the filament suppresses the peaks of mass numbers 24 and 38.

REFERENCES

- Arnot, F. L. 1937*a* *Proc. Roy. Soc. A*, **158**, 137.
— 1937*b* *Proc. Roy. Soc. A*, **158**, 157.
Arnot, F. L. and Beckett, C. 1938 *Nature, Lond.*, **141**, 1011.

- Arnot, F. L. and Milligan, J. C. 1936*a* *Proc. Roy. Soc. A*, **153**, 359.
 — — 1936*b* *Proc. Roy. Soc. A*, **156**, 538.
 Bleakney, W. 1930 *Phys. Rev.* **36**, 1303.
 Bradbury, N. E. 1933 *Phys. Rev.* **44**, 883.
 Loeb, L. B. 1935 *Phys. Rev.* **48**, 684.
 Lozier, W. W. 1932 *Phys. Rev.* **42**, 518.
 Mann, W. B. 1937 *Proc. Roy. Soc. A*, **161**, 236.
 Massey, H. S. W. 1937 *Proc. Roy. Soc. A*, **163**, 542.
 Reimann, A. L. 1935 *Phil. Mag.* **20**, 594.
 Sloane, R. H. and Press, R. 1938 *Nature, Lond.*, **141**, 872.
 Smith, R. A. 1938 *Proc. Roy. Soc. A*, **168**, 19.
 Smithells, C. J. and Ransley, C. E. 1936 *Proc. Roy. Soc. A*, **155**, 195.

New systems of normal co-ordinates for relativistic optics

BY G. TEMPLE

(*Communicated by Sir Arthur Eddington, F.R.S.—Received 22 June 1938*)

INTRODUCTION

The object of this paper is to introduce into the general theory of relativity some new systems of normal co-ordinates which are especially adapted to the discussion of problems of astronomical optics. In the classical system of normal co-ordinates introduced by Riemann the geodesics which start from a fixed base point in a metric manifold are represented by linear equations; in the new systems of normal co-ordinates null geodesics from points on a fixed base curve are represented by linear equations or form a system of parametric curves. If the system of null geodesics consists of the "forward" null geodesics starting from points on the fixed curve it represents the rays of light emitted by a source whose world line is the prescribed curve. If the system of null geodesics consists of the "backward" null geodesics ending at points on the fixed curve it represents the rays of light received by an observer whose world line is the given curve. It therefore seems appropriate to describe the new systems of co-ordinates as "optical co-ordinates".

The use of optical co-ordinates considerably simplifies general optical theory in relativistic form. Fermat's Principle can be rigorously established and the treatment of the various astronomical determinations of distance

can be put in a simple and concise form. Some previous discussions of these topics have been open to question in view of their unjustified use of Riemannian normal co-ordinates on the null cone of the base point. The general results obtained are applied to the case of greatest practical interest—the isotropic expanding universe.

1. NULL GEODESICS

The general theory of this paper applies to a real n -dimensional manifold M with a metric specified by an *indefinite* quadratic form,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu,$$

whose signature is $+1 - (n-1)$. In such a manifold the length of an arc of a real curve may be zero, imaginary or even complex, and accordingly the arc-length of a curve is not suitable for use as a parameter. We therefore write the equations of any real curve C in M in the general form

$$x^\mu = x^\mu(\sigma),$$

where σ is an arbitrary real parameter and the functions $x^\mu(\sigma)$ are real functions. We indicate differentiation with respect to σ by dots.

The function
$$L = g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \tag{1.1}$$

is not an invariant of C at any specified point P , for its value at P depends upon the choice of the parameter σ ; but it is obvious that the function

$$\text{sgn } L = \begin{cases} L/|L| & \text{if } L \neq 0, \\ 0 & \text{if } L = 0, \end{cases}$$

is an invariant of C at any point P , and we shall describe the direction of C at P as time-like, space-like or null, accordingly as $\text{sgn } L = +1, -1$, or 0 .

We shall be especially interested in geodesics whose direction is everywhere null, and we therefore adopt a slight extension of the usual definition, which is applicable only to geodesics whose direction is everywhere time-like or everywhere space-like. We first define a “transport vector”^{*} for a curve C as a vector K^μ attached to each point of C in such a way that the vectors at any two points are parallel (in the sense of Levi-Civita) for displacements along C . We then define a geodesic to be a curve G such that there exists a

^{*} This useful term was introduced in a paper by Kermack, McCrea and Whittaker (1933) and is used here in a slightly wider sense. The definition of a geodesic adopted here is an obvious generalization of that given by these authors.

transport vector $K^\mu(\sigma)$ which is *parallel* to the direction of G at each point σ . These properties of a geodesic are expressed by the two equations

$$\left. \begin{aligned} \dot{K}^\mu &= -\Gamma_{\alpha\beta}^\mu K^\alpha \dot{x}^\beta, \\ \text{and} \quad K^\mu &= \phi(\sigma) \dot{x}^\mu, \end{aligned} \right\} \quad (1.2)$$

where $\phi(\sigma)$ is an undetermined function of σ which is always positive. It follows that $x^\mu(\sigma)$ satisfies the equation

$$\ddot{x}^\mu + \Gamma_{\alpha\beta}^\mu \dot{x}^\alpha \dot{x}^\beta = -\frac{d \log \phi(\sigma)}{d\sigma} \dot{x}^\mu. \quad (1.3)$$

It is clear from the theory of parallel displacement that the function

$$g_{\mu\nu} K^\mu K^\nu = [\phi(\sigma)]^2 L \quad (1.4)$$

has a constant value along a geodesic G , so that $\text{sgn } L$ has the same value at all points of G . Hence the type of a geodesic, viz. time-like, space-like or null, is the same throughout its length, and we can therefore classify geodesics as time-like, space-like or null.

For all types of geodesics we can define a new parameter λ (apart from an additive constant) by the equation

$$\frac{d\lambda}{d\sigma} = \frac{1}{\phi(\sigma)}, \quad (1.5)$$

when equations (1.2) become

$$\begin{aligned} \frac{dK^\mu}{d\lambda} &= -\Gamma_{\alpha\beta}^\mu K^\alpha \frac{dx^\beta}{d\lambda}, \\ \text{and} \quad K^\mu &= \frac{dx^\mu}{d\lambda}. \end{aligned}$$

These relations yield the equations to a geodesic in the standard form

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0, \quad (1.6)$$

which is seen to be valid whether the geodesic is time-like, space-like or null.

To determine the significance of the new parameter λ we introduce a second transport vector A^μ for the geodesic G , with the condition that the length of A^μ is to be +1 or -1 accordingly as A^μ is time-like or space-like. It then follows from the theory of parallel displacement that $A_\mu K^\mu$ or $A_\mu dx^\mu/d\lambda$ has a constant value, say k , along G , whence

$$\int_{P_0} A_\mu dx^\mu = \int_{P_0} k d\lambda = k(\lambda - \lambda_0). \quad (1.7)$$

Now $A_\mu dx^\mu$ is the projection of the element of the geodesic between the points (x^μ) and $(x^\mu + dx^\mu)$ upon the transport vector A^μ . We shall therefore call $\int_{P_0} A_\mu dx^\mu$ the "projected length" of the geodesic* from P_0 with respect to the "projection vector" A_μ .

It is now clear that the value of the parameter λ at a point P is merely a linear function of the projected length of the geodesic arc P_0P . The same equation also shows that a different choice of the transport vector A_μ changes all projected lengths along G in the same ratio. For, if B_μ is another unit transport vector along G , it follows from (1.7) that

$$\int_{P_0} A_\mu dx^\mu : \int_{P_0} B_\mu dx^\mu = \left(A_\mu \frac{dx^\mu}{d\lambda} \right)_P : \left(B_\mu \frac{dx^\mu}{d\lambda} \right)_P. \quad (1.8)$$

Hence the projected length with respect to A_μ is proportional to the projection of A_μ on G at any point.

2. PROJECTED NORMAL CO-ORDINATES

The system of normal co-ordinates introduced by Riemann is defined in terms of a set of n orthogonal directions at the base point O . The normal co-ordinates of a point P at a geodesic distance s from O are taken to be

$$z^\mu = l^\mu s,$$

where (l^1, l^2, \dots, l^n) are the direction cosines of the geodesic at O with respect to the n orthogonal directions. In a manifold with an indefinite metric this definition does not apply to points P on null geodesics from O . We can, however, extend the definition so as to apply to points on all types of geodesic from O by employing the concept of projected length.

We therefore take the normal co-ordinate z^μ of P to be the projected length of the geodesic OP with respect to a projection vector in the μ th orthogonal direction at O . It is clear that the directions of the co-ordinate axes at O will coincide with the n orthogonal directions. Hence, if

$$ds^2 = g_{\mu\nu} dz^\mu dz^\nu,$$

$$\text{at the point } O, \quad g_{\mu\nu} = 0 \quad \text{when } \mu \neq \nu. \quad (2.1)$$

We can also show that $g_{\mu\mu} = e_\mu$, the indicator of the μ th direction, i.e. a scalar with the value $+1$ if the direction is time-like and the value -1 if

* The concept of "projected length" was first introduced, for null geodesics, by Ruse (1933).

the direction is space-like. For if A^ρ is the projection vector for the μ th direction, at the point O

$$A^\rho = \delta_\mu^\rho (e_\mu / g_{\mu\mu})_0^\frac{1}{2}, \quad (\text{not summed for } \mu).$$

Hence, if P is near O , it follows from the definition of projected distance that

$$dz^\mu = g_{\rho\sigma} A^\rho dz^\sigma = g_{\mu\mu} (e_\mu / g_{\mu\mu})^\frac{1}{2} dz^\mu, \quad (\text{not summed for } \mu).$$

Therefore, at O , (2.2)

$$g_{\mu\mu} = e_\mu.$$

Hence at O the projection vector for the μ th direction is $A^\rho = \delta_\mu^\rho$. Now, let λ be the projected length of the geodesic OP with respect to a direction at O specified by a unit vector t^μ . Then, by (1.8),

$$z^\mu \propto (g_{\rho\sigma} A^\rho dz^\sigma / d\lambda)_0 = e_\mu (dz^\mu / d\lambda)_0, \quad (\text{not summed for } \mu),$$

and (2.3)

$$\lambda \propto (g_{\rho\sigma} t^\rho dz^\sigma / d\lambda)_0 = (e_\rho t^\rho dz^\rho / d\lambda)_0.$$

Therefore, if

$$a^\mu = e_\mu (dz^\mu / d\lambda)_0 \div (e_\rho t^\rho dz^\rho / d\lambda)_0, \quad (\text{not summed for } \mu), \quad (2.3)$$

we can write the equations of the geodesic OP in the parametric form,

$$z^\mu = a^\mu \lambda, \quad (2.4)$$

which will be valid for all geodesics from O except those whose initial direction is orthogonal to t^μ . If t^μ is time-like this parametric representation is valid for all time-like and null geodesics from O , and it is also valid for almost all space-like geodesics from O . It is this parametric representation which provides the real justification for the methods employed by Etherington (1933) and Ruse (1933). We notice that

$$(t_\mu a^\mu)_0 = t^\mu a_\mu = 1, \quad (2.5)$$

which is the relation employed by Etherington (equation (13)).

If we take t^μ to be along a time-like geodesic OP , then λ becomes the arc length s , and $t^\mu = (dz^\mu / ds)_0$. Hence, if

$$b^\mu = e_\mu (dz^\mu / d\lambda)_0, \quad (\text{not summed for } \mu),$$

the time-like geodesics can be represented parametrically as

$$z^\mu = b^\mu s.$$

This result establishes the identity of our "projected normal co-ordinates" (z^μ) with Riemannian normal co-ordinates.

The vector with components $(dz^\mu/d\lambda)$ is clearly a transport vector along the geodesic $z^\mu = a^\mu\lambda$, and its length is therefore a constant. Hence

$$\left(\frac{ds}{d\lambda}\right)_P^2 = \left\{g_{\mu\nu} \frac{dz^\mu}{d\lambda} \frac{dz^\nu}{d\lambda}\right\}_P = \{g_{\mu\nu} a^\mu a^\nu\}_0 = (a_\mu a^\mu)_0.$$

therefore the geodesic distance OP is

$$s_P = \lambda_P (a_\mu a^\mu)_0^{\frac{1}{2}} = (z_\mu z^\mu)^{\frac{1}{2}},$$

and the characteristic function for O and P is

$$\Omega = \frac{1}{2}s_P^2 = \frac{1}{2}z_\mu z^\mu. \quad (2.6)$$

Now let the point P describe a curve whose direction is parallel to t^μ at O (with respect to displacement along the geodesic OP). Then, if τ is the arc length along this curve,

$$\frac{d\Omega}{d\tau} = z_\mu \frac{dz^\mu}{d\tau} = \lambda a_\mu \frac{dz^\mu}{d\tau},$$

and, by the theory of parallel displacement,

$$\left(a_\mu \frac{dz^\mu}{d\tau}\right)_P = (a_\mu t^\mu)_0 = 1 \quad \text{by (2.5).}$$

Hence
$$\frac{d\Omega}{d\tau} = \lambda, \quad (2.7)$$

i.e. the rate of change of $\Omega(O, P)$ with respect to the arc length of the curve C described by P is equal to the projected length of the geodesic OP , the projection vector being the unit tangent to C at P . This result is valid whether the geodesic OP is time-like, space-like or null. In the last case the theorem has been given by Ruse (1933, § 4).

3. OPTICAL CO-ORDINATES

In a manifold of signature $+1-(n-1)$, of any n orthogonal directions at a point O , one is time-like and the other $(n-1)$ are space-like. Hence, by (2.2) in the neighbourhood of O , the metric in projected normal co-ordinates is given by

$$ds^2 = (dz^n)^2 - (dz^1)^2 - \dots - (dz^{n-1})^2,$$

taking the n th direction to be time-like with indicator $e_n = 1$, and the other directions to be space-like with indicators $e_j = -1$. The null directions at O then consist of two distinct systems, (i) the "forward" null directions for which $dz^n > 0$ and (ii) the "backward" null directions for which $dz^n < 0$.

This distinction is independent of the choice of the time-like direction specified by $dz^1=0$, $dz^2=0$, ..., $dz^{n-1}=0$ when the signature is $+1-(n-1)$ (Hadamard 1923, p. 39).

We can now define the two systems of optical co-ordinates in terms of the forward and backward null geodesics at points on an arbitrarily chosen time-like curve C . Let A_0 be any fixed point on the base curve C and let A be any other point on C . Let the arc-length A_0A along C be denoted by τ . Each of the forward null geodesics starting from A_0 is specified by the $(n-1)$ ratios of the n displacements dz^μ along the null geodesic at A_0 . These $(n-1)$ ratios are connected by the relation $g_{\mu\nu}dx^\mu dx^\nu = 0$. Hence each null geodesic can be specified by $(n-2)$ independent constants, a^1, a^2, \dots, a^{n-2} . The equations to the forward null geodesics through A_0 can therefore be written in the form

$$x^\mu = f^\mu(\lambda; a^1, a^2, \dots, a^{n-2}),$$

λ being the projected length along the null geodesics (measured from A_0) with respect to the direction of C at A_0 .

We next consider the forward null geodesics drawn from any other point A of C . Each of these is specified by $(n-2)$ independent constants, and we shall choose these constants so that null geodesics at A_0 and A with the same set of constants have initial directions which are parallel with respect to displacement along C . The equations to the complete set of forward null geodesics drawn from all points on C can then be written in the form

$$x^\mu = f^\mu(\lambda; a^1, a^2, \dots, a^{n-2}, \tau).$$

Here λ is the projected length from A along the geodesic through A with respect to the direction of C at A , and τ is the arc-length A_0A along C .

There will exist a domain D of the manifold M in the neighbourhood of C such that no two of the forward null geodesics drawn from points of C intersect in D (except on C itself). In this domain we adopt the co-ordinates (y^μ) defined by the equations

$$\left. \begin{aligned} y^p &= a^p, & (p = 1, 2, \dots, n-2), \\ y^{n-1} &= \tau, \\ y^n &= \lambda. \end{aligned} \right\} \quad (3.1)$$

These form the forward system of optical co-ordinates based on the curve C . The backward system of optical co-ordinates is defined similarly in terms of the backward null geodesics at points of C , and the following theory applies, *mutatis mutandis*, to both systems.

We proceed to determine the particularly simple form assumed by the metric in optical co-ordinates. Let

$$ds^2 = h_{\mu\nu} dy^\mu dy^\nu.$$

It will be convenient to allow the Greek suffixes to run from 1 to n , while i, j, k run from 1 to $n-1$, and p, q from 1 to $n-2$.

Since C is given by the single equation, $\lambda = 0$, on considering a displacement along C it follows that $h_{jk} dy^j dy^k = d\tau^2$ if $\lambda = 0$, i.e.

$$\left. \begin{aligned} h_{n-1, n-1} &= 1 \\ h_{pq} &= 0, \quad h_{p, n-1} = 0 \end{aligned} \right\} \quad \text{if } \lambda = 0. \quad (3.2)$$

Again, by considering a displacement along a null geodesic at A ,

$$h_{\mu\nu} dy^\mu dy^\nu = h_{nn} d\lambda^2, \quad \text{if } \lambda = 0,$$

$$\text{i.e.} \quad h_{pn} = 0 \quad \text{if } \lambda = 0. \quad (3.3)$$

The equations of the null geodesics from A , viz. $y^j = \text{constant}$, $y^n = \lambda$, must satisfy the equations

$$\frac{d^2 y^\mu}{d\lambda^2} + \Gamma_{\alpha\beta}^\mu \frac{dy^\alpha}{d\lambda} \frac{dy^\beta}{d\lambda} = 0$$

$$\text{and} \quad h_{\mu\nu} dy^\mu dy^\nu = 0.$$

$$\text{It follows that} \quad \Gamma_{nn}^\mu = 0 \quad \text{and} \quad h_{nn} = 0 \quad \text{everywhere in } D. \quad (3.4)$$

$$\text{Hence} \quad 2 \frac{\partial h_{\mu n}}{\partial \lambda} - \frac{\partial h_{nn}}{\partial y^\mu} = 2 \Gamma_{nn, \mu} = 0,$$

i.e. by (3.4), $\partial h_{\mu n} / \partial \lambda = 0$ everywhere in D . Therefore

$$\left. \begin{aligned} h_{n-1, n} &\text{ is independent of } \lambda, \\ \text{and, by (3.3), } h_{pn} &= 0 \text{ everywhere.} \end{aligned} \right\} \quad (3.5)$$

Now consider the projection vector A^μ along a null geodesic, $y^j = \text{constant}$. At A , $A^\mu = \delta_{n-1}^\mu$, and along C

$$\lambda = \int_A h_{\mu\nu} A^\mu dy^\nu = \int_0 h_{\mu n} A^\mu d\lambda,$$

$$\text{i.e.} \quad h_{\mu n} A^\mu = 1. \quad (3.6)$$

$$\text{Hence, at } A, \quad 1 = h_{\mu n} \delta_{n-1}^\mu = h_{n, n-1}.$$

$$\text{Therefore, by (3.5),} \quad h_{n, n-1} = 1 \text{ everywhere.} \quad (3.7)$$

Now, by (3.4), (3.5) and (3.7),

$$h_{\mu n} A^\mu = A^{n-1},$$

whence

$$A^{n-1} = 1. \quad (3.8)$$

Since the initial directions of null geodesics with the same parameters, a^1, a^2, \dots, a^{n-2} , are parallel with respect to displacements along C , the vector K^μ with components $K^\mu = \delta_n^\mu$ must satisfy the equation

$$\frac{dK^\mu}{ds} + \Gamma_{\alpha\beta}^\mu K^\alpha \frac{dy^\beta}{ds} = 0$$

whenever λ is zero. Hence we must have $\Gamma_{nj}^\mu = 0$ if $\lambda = 0$, i.e.

$$\frac{\partial h_{j\mu}}{\partial \lambda} + \frac{\partial h_{n\mu}}{\partial y^j} - \frac{\partial h_{nj}}{\partial y^\mu} \equiv 2\Gamma_{nj,\mu} = 0, \quad \text{if } \lambda = 0.$$

Now, by (3.5), (3.7) and (3.4),

$$h_{np} = 0, \quad h_{n,n-1} = 1 \quad \text{and} \quad h_{nn} = 0 \quad \text{everywhere.}$$

Therefore

$$\partial h_{j\mu} / \partial \lambda = 0 \quad \text{if } \lambda = 0.$$

Hence, by (3.2) and (3.3) we can write

$$\left. \begin{aligned} h_{jk} &= \lambda^2 \eta_{jk} \quad \text{unless } j = n-1, k = n-1, \\ \text{and} \quad h_{n-1,n-1} &= 1 + \lambda^2 \eta_{n-1,n-1}, \end{aligned} \right\} \quad (3.9)$$

where the coefficients η_{jk} are regular functions of λ in D . The expression for the metric then assumes the simplified form

$$ds^2 = \lambda^2 \eta_{jk} dy^j dy^k + (dy^{n-1})^2 + 2dy^{n-1} dy^n.$$

We note for future reference that

$$\begin{aligned} h \equiv \det h_{\mu\nu} &= -\det h_{pq} = -\lambda^{2n-4} \det \eta_{pq} \\ &= -\lambda^{2n-4} \eta. \end{aligned} \quad (3.10)$$

In the reciprocal tensor $h^{\mu\nu}$,

$$h^{n-1,j} = 0 \quad \text{and} \quad h^{n-1,n} = 1. \quad (3.11)$$

By direct calculation we find that on the curve C , the components of the Riemann-Christoffel tensor $R_{\mu\nu,\rho\sigma}$ have the following values:

$$R_{ij,kl} = O(\lambda^2),$$

$$R_{ij,kn} = O(\lambda),$$

and

$$R_{jn,kn} = 2\eta_{jk}.$$

4. LIGHT RAYS AND WAVE FRONTS

Although the problem of the propagation of *continuous* waves in a manifold with an arbitrarily prescribed metric is of considerable complexity (Hadamard 1923), the theory of isolated, discontinuous waves or *signals* is peculiarly simple. In this theory the wave fronts and light rays are considered primarily as loci, say W and R , of three and one dimensions respectively in the four-dimensional space-time manifold. To obtain the physical interpretation of the theoretical results we have to consider the intersections of these loci with the instantaneous space S of a moving observer. We thus obtain a moving system of two dimensional wave fronts W^* in S and a moving point which traces out the light track R^* in S .

The fundamental theorems (Hadamard 1903) are (1) if $\phi(x^1, x^2, x^3, x^4) = 0$ is the equation of a wave front W in S , then ϕ satisfies the equation

$$\Delta\phi \equiv g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} = 0, \quad (4.1)$$

where

$$\phi_{,\mu} = \partial\phi/\partial x^{\mu};$$

and (2) the rays R are the orthogonal trajectories of the wave front W . From these principles it follows that the rays R lie wholly in a wave front W , and that they are null geodesics.

The system of optical co-ordinates provides a simple illustration of these results. Since

$$h^{\mu\nu} \frac{\partial y^3}{\partial y^{\mu}} \frac{\partial y^3}{\partial y^{\nu}} = h^{33} = 0, \quad \text{by (3.11),}$$

if $n=4$, it follows that the surfaces $\tau = \text{constant}$ are wave fronts. The orthogonal trajectories of any such wave front are given by

$$\frac{dy^{\mu}}{d\sigma} = f(\sigma) h^{\mu\nu} \frac{\partial \tau}{\partial y^{\nu}} = f(\sigma) h^{\mu 3} = f(\sigma) \delta^{\mu}_4, \quad \text{by (3.11).}$$

Hence the orthogonal trajectories are the null geodesics $y^j = \text{constant}$, and the wave fronts $\tau = \text{constant}$ therefore correspond to signals emitted by a source whose world line is the basic curve C .

This result suggests a generalisation of the system of optical co-ordinates which is appropriate for the discussion of any system of wave fronts. Returning to n dimensions let $\phi(x^1, x^2, \dots, x^n) = \text{constant}$ satisfy the equation

$$\Delta\phi \equiv g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} = 0. \quad (4.2)$$

Any systems of surfaces $f(x^1, x^2, \dots, x^n) = \text{constant}$, which are orthogonal to the system $\phi = \text{constant}$, satisfy the equation

$$g^{\mu\nu} \frac{\partial f}{\partial x^\mu} \frac{\partial \phi}{\partial x^\nu} = 0. \quad (4.3)$$

This equation possesses $(n-1)$ independent integrals, of which one is clearly $f = \phi$ by (4.2). Let $(n-2)$ other independent integrals be given by $f^p = \text{constant}$ ($p = 1, 2, \dots, n-2$). Then the equations $f^p = \text{constant}$, $\phi = \text{constant}$, specify a system of curves R which are the orthogonal trajectories of the surfaces (W) $\phi = \text{constant}$. These curves are therefore null geodesics.

To complete the co-ordinate system we introduce finally a system of surfaces, $\lambda = \text{constant}$, such that

$$g^{\mu\nu} \frac{\partial \lambda}{\partial x^\mu} \frac{\partial \phi}{\partial x^\nu} = 1, \quad (4.4)$$

and we write the equations of the null geodesics R in the form

$$x^\mu = x^\mu(\lambda).$$

Since these curves are orthogonal to the surfaces W ,

$$\frac{dx^\mu}{d\lambda} = f(\lambda) g^{\mu\nu} \frac{\partial \phi}{\partial x^\nu},$$

whence

$$f(\lambda) g^{\mu\nu} \frac{\partial \lambda}{\partial x^\mu} \frac{\partial \phi}{\partial x^\nu} = \frac{dx^\mu}{d\lambda} \frac{\partial \lambda}{\partial x^\mu} = 1.$$

Therefore, by (4.4), $f(\lambda) = 1$, and

$$\frac{dx^\mu}{d\lambda} = \phi^\mu.$$

Hence

$$\frac{d^2 x^\mu}{d\lambda^2} = \frac{dx^\nu}{d\lambda} D_\nu \phi^\mu - \Gamma_{\nu\alpha}^\mu \phi^\alpha \frac{dx^\nu}{d\lambda},$$

where D_ν is the symbol for covariant differentiation. Then

$$\begin{aligned} \frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\nu\alpha}^\mu \frac{dx^\nu}{d\lambda} \frac{dx^\alpha}{d\lambda} &= \phi^\nu D_\nu \phi^\mu = g^{\mu\rho} \phi^\nu D_\nu \phi_\rho \\ &= g^{\mu\rho} \phi^\nu D_\rho \phi_\nu \\ &= \frac{1}{2} g^{\mu\rho} \frac{\partial}{\partial x^\rho} (\phi^\nu \phi_\nu) = 0, \quad \text{by (4.2).} \end{aligned}$$

Therefore λ is a projected length measured along the null geodesics R .

Our generalized system of optical co-ordinates is therefore

$$y^p = f^p, \quad y^{n-1} = \phi, \quad y^n = \lambda. \quad (4.5)$$

If the expression for the metric is

$$ds^2 = h_{\mu\nu} dy^\mu dy^\nu,$$

and $h^{\mu\nu}$ is the tensor reciprocal to $h_{\mu\nu}$, then

$$h^{n-1, n-1} = g^{\mu\nu} \frac{\partial y^{n-1}}{\partial x^\mu} \frac{\partial y^{n-1}}{\partial x^\nu} = 0 \quad \text{by (4.2),}$$

$$h^{p, n-1} = g^{\mu\nu} \frac{\partial y^p}{\partial x^\mu} \frac{\partial y^{n-1}}{\partial x^\nu} = 0 \quad \text{by (4.3),}$$

$$\text{and} \quad h^{n, n-1} = g^{\mu\nu} \frac{\partial y^n}{\partial x^\mu} \frac{\partial y^{n-1}}{\partial x^\nu} = 1 \quad \text{by (4.4).}$$

It then follows that

$$h_{pn} = 0, \quad h_{n-1, n} = 1 \quad \text{and} \quad h_{nn} = 0, \quad (4.6)$$

so that the metric has the same form as in the system of optical co-ordinates introduced in § 3.

5. VARIATIONAL PROPERTIES OF NULL GEODESICS

The discussion of the variational properties of null geodesics turns upon the fact that on the null cone of any point the metric is negative semi-definite. To prove this we introduce the three vectors, A^μ , B^μ and C^μ , which are specified at any point in optical co-ordinates by the conditions:

$$(i) \quad A^\mu = \delta_{h-1}^\mu, \quad (ii) \quad B^p = 0, \quad B^{n-1} = 1, \quad B^n = -h_{n-1, n-1},$$

$$(iii) \quad C^p \neq 0, \quad C^{n-1} = 0, \quad C^n = -h_{n-1, p} C^p.$$

It is easily verified that these vectors are mutually orthogonal. Now, in a manifold of signature $+1-(n-1)$, any vector orthogonal to a time-like vector is space-like. But, if $h_{n-1, n-1} > 0$, A^μ is time-like, and so C^μ is space-like. And, if $h_{n-1, n-1} < 0$, B^μ is time-like, and again C^μ is space-like. Hence, in both cases,

$$h_{jk} C^j C^k + 2C^{n-1} C^n < 0,$$

i.e.

$$h_{pq} C^p C^q < 0.$$

Now the components C^1, C^2, \dots, C^{n-2} of C^μ can be chosen arbitrarily. Hence if (dy^μ) is any displacement on the null cone of the point A ,

$$h_{pq} dy^p dy^q < 0, \quad (5.1)$$

unless the displacement is along a null geodesic, when $h_{pq} dy^p dy^q = 0$. Therefore the quadratic form $h_{pq} dy^p dy^q$ is negative semi-definite.

We now consider the first variation of the integral $\int ds$ for null geodesics. We compare an arc of a null geodesic G ,

$$y^j = \text{a constant}, \quad a^j, \quad 0 \leq \lambda \leq A,$$

and an arc of a neighbouring curve H ,

$$y^j = a^j + \epsilon \eta^j(\lambda), \quad 0 \leq \lambda \leq A,$$

with the same end-points, so that

$$\eta^j(0) = 0, \quad \eta^j(A) = 0.$$

The value of $\int ds$ taken along H is

$$J(\epsilon) = \int_0^A \{2\epsilon \dot{\eta}^{n-1} + \epsilon^2 h_{jk} \dot{\eta}^j \dot{\eta}^k\}^{\frac{1}{2}} d\lambda,$$

where dots denote differentiation with respect to λ .

If $\dot{\eta}^{n-1}$ is not zero all along H , then

$$J(\epsilon)/\epsilon^{\frac{1}{2}} \rightarrow \int_0^A (2\dot{\eta}^{n-1})^{\frac{1}{2}} d\lambda, \quad \text{as } \epsilon \rightarrow 0.$$

Now the first variation of $J(\epsilon)$ is $\lim_{\epsilon \rightarrow 0} J(\epsilon)/\epsilon$. Hence in this case the first variation is infinite.

If $\dot{\eta}^{n-1}$ is zero all along H , the variation vector $\epsilon \eta^j$ is orthogonal to the null geodesic considered, and

$$J(\epsilon)/\epsilon = \int_0^A (\eta_{pq} \dot{\eta}^p \dot{\eta}^q)^{\frac{1}{2}} d\lambda.$$

But, by (5.1), the quadratic form $h_{pq} \dot{\eta}^p \dot{\eta}^q$ is essentially negative. Hence the first variation of $J(\epsilon)$ is a non-vanishing, wholly imaginary number. This property of null geodesics is in sharp contrast with the corresponding property of time-like or space-like geodesics, for which the first variation of $\int ds$ is always zero.

In contrast with this negative result we can, however, obtain a positive variational property of null geodesics by generalizing Fermat's Principle. The classical form of this principle is that the path in space of a ray of light is such as to make the time of passage minimum, i.e. if v is the speed of light at any point and dl is an element of a space curve joining two points A , B , then the integral $\int_A^B dl/v$ is a minimum along the actual path of the ray, if A and B are sufficiently near together. To express this principle in relativistic form we must consider two signals emitted instantaneously from the same point A , the world line of the first being a null geodesic and the world line of the second being any null curve. Then both signals move with the speed of light at each point of their world lines. Let the world lines of these two signals intersect the world line of an observer at points P and Q . The times of passage of the signals are not directly observed, but the order of arrival of the signals is directly observable and depends upon the order of the points P , Q on the observer's world line. If the variation of the null curve from the null geodesic is small, we shall show (i) that P and Q coincide to the first order of small quantities and (ii) that to the next order of approximation Q is later than P in the experience of the observer.

In proving these results we take A to be on the basic curve C of our optical co-ordinates. At A , $h_{n-1, n-1} = 1$. Hence in some sub-domain of D , say D' , $h_{n-1, n-1}$ is positive. We shall consider only the case in which P and Q lie in D' .

Let $y^j = a^j$ be the equations of the null geodesic, and let

$$y^j = a^j + \epsilon \eta^j(\lambda), \quad \eta^j(0) = 0,$$

be the equations of the null curve. Then

$$\epsilon^2 h_{pq} \dot{\eta}^p \dot{\eta}^q + 2\epsilon^2 h_{p, n-1} \dot{\eta}^p \dot{\eta}^{n-1} + \epsilon^2 h_{n-1, n-1} (\dot{\eta}^{n-1})^2 + 2\epsilon \dot{\eta}^{n-1} = 0.$$

Solving this quadratic for $\dot{\eta}^{n-1}$ we find that

$$\dot{\eta}^{n-1} = -\frac{1}{2}\epsilon h_{pq} \dot{\eta}^p \dot{\eta}^q + O(\epsilon^2).$$

Therefore

$$\left. \begin{aligned} \eta^{n-1}(\lambda) &= \frac{1}{2}\epsilon I + O(\epsilon^2), \\ \text{where } I &= -\int_0^\lambda h_{pq} \dot{\eta}^p \dot{\eta}^q d\lambda, \end{aligned} \right\} \quad (5.2)$$

and, by (5.1), $I \geq 0$.

There is no loss of generality in taking the world line of the observer to be

$y^p = a^p$, $y^n = a^n$, for we are free to choose the direction of C at A arbitrarily, provided it remains time-like. Then

$$(y^{n-1})_P = a^{n-1}$$

$$\text{and} \quad (y^{n-1})_Q = a^{n-1} + \epsilon \eta^{n-1}(\lambda) = a^{n-1} + \frac{1}{2} \epsilon^2 I + O(\epsilon^3). \quad (5.3)$$

It follows from (5.2) that to the first order of approximation P and Q coincide, and to the second order, $(y^{n-1})_Q > (y^{n-1})_P$. Now, if N^μ is a vector along the forward null geodesic, $N^i = 0$ and $N^n > 0$; and, if T^μ is a unit vector in the direction of y^{n-1} increasing, $T^p = 0$, $T^{n-1} = (h_{n-1, n-1})^{-\frac{1}{2}}$, $T^n = 0$. Hence the projection of N^μ on T^μ is

$$h_{jk} N^j T^k + N^{n-1} T^n + N^n T^{n-1} = N^n (h_{n-1, n-1})^{-\frac{1}{2}}.$$

Since this is positive the direction of y^{n-1} increasing is the forward direction along the observer's world line. Therefore, by (5.3), Q is later than P . This result is the relativistic generalization of Fermat's Principle.

6. ASTRONOMICAL DISTANCES

We now proceed to apply the general theory of optical co-ordinates to the manifold of general relativity with a view to obtaining explicit expressions for the different types of "distance" specified by astronomical observations. There is no restriction, in this section, on the kind of manifold, but in the next section we indicate the particular results applicable to the manifold of an expanding universe of Lemaître's type.

(a) *Projected distance*

In the astronomical problem we have to discuss the "distance" between a star S at the instant of the emission of a signal and an observer P at the instant of the reception of this signal. S and P are therefore on the same null geodesic, and the preceding theory shows that the parameter which is fundamental for the geometry of null geodesics is the projected distance. This parameter was introduced into the astronomical problem by Kermack, McCrea and Whittaker (1933), whose definition of "spatial distance" is equivalent to what we have described above as the projected length of the null geodesic SP when the projection vector is along the observer's world line at P . Ruse (1933) has shown that the "spatial distance" so defined has a physical interpretation as the distance SP measured by rigid rods by a line of observers on the null geodesic SP .

When S is near to P we can identify the spatial distance so defined by taking orthogonal Galilean co-ordinates t, x, y, z at P with the time-like

t -axis along P 's world line. (These are in fact projected normal co-ordinates at P .) Then the metric is

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2.$$

The forward null geodesic from the star at (t_S, x_S, y_S, z_S) will pass through P at $(0, 0, 0, 0)$ if

$$t_S = -(x_S^2 + y_S^2 + z_S^2)^{\frac{1}{2}},$$

and its projected length (in the direction of the t -axis at P) will be

$$\int_{t_S}^0 dt = -t_S = (x_S^2 + y_S^2 + z_S^2)^{\frac{1}{2}}. \quad (6.1)$$

Hence when S is near P the "spatial distance" reduces to the ordinary distance measured in the instantaneous space of the observer.

(b) Doppler shift

There is clearly an alternative definition of spatial distance as the projected length of the null geodesic SP when the projection vector is along the world line of the star at S . Let λ' and λ'' denote the projected lengths when the projection vector is along the world line of the observer and star respectively. Then we shall show that the ratio $\lambda'' : \lambda'$ is the familiar Doppler factor giving the ratio of the observed to the emitted wave-lengths. If we adopt optical co-ordinates with the world line of S as the base curve, and if the world line of P is $y^\mu = y_P^\mu(s)$, s being the arc length, then it follows from (1.8) that

$$\lambda' : \lambda'' = h_{\mu\nu} \frac{dy_P^\mu}{ds} \frac{dy_G^\nu}{d\lambda} : h_{\alpha\beta} \frac{dy_S^\alpha}{ds} \frac{dy_G^\beta}{d\lambda},$$

$y^\nu = y_G^\nu(\lambda)$ being the equations of the null geodesic SP and $y^\rho = y_S^\rho(s)$ the equations of the base curve. Hence

$$\lambda' : \lambda'' = dy_G^3/ds. \quad (6.2)$$

From the standpoint of optics the Doppler effect for light signals can be discussed either in terms of the energy momentum vector of photons or of the frequency of a series of signals. The energy momentum vector E^μ of a photon is a transport vector parallel at every point to the world line R of a photon. Hence, in optical co-ordinates,

$$E^j = 0 \quad \text{and} \quad E^4 = E,$$

E being a constant along R . The energy of the photon, as measured by an observer P , is the component of E^μ along the world line of P , i.e. it is

$$h_{\mu\nu} E^\mu dy_P^\nu/ds = E dy_P^3/ds.$$

For an observer moving along the base curve C , $dy_S^3 = ds$, so that such an observer estimates the energy as E . Now the Doppler shift for photons is

$$D = \frac{\text{wave-length at } P}{\text{wave-length at } S} = \frac{\text{frequency at } S}{\text{frequency at } P} = \frac{\text{energy at } S}{\text{energy at } P}.$$

Hence

$$D = ds/dy_P^3, \quad (6.3)$$

and therefore

$$D = \lambda''/\lambda'.$$

Again if two signals are emitted from points A and B on the base curve C with co-ordinates $(a, b, \tau, 0)$ and $(a+da, b+db, \tau+d\tau, 0)$, and are received by an observer at points P and Q with co-ordinates (a, b, τ, λ) and $(a+da, b+db, \tau+d\tau, \lambda+d\lambda)$, then the effective frequencies of emission and reception are proportional to $(d\tau)^{-1}$ and $(ds)^{-1}$ respectively, where ds is the interval PQ . Hence

$$\begin{aligned} D &= (d\tau)^{-1}/(ds)^{-1} = ds/d\tau \\ &= ds/dy_P^3. \end{aligned}$$

Thus the same value for the Doppler shift is given by the corpuscular and undulatory theories, as was first shown by Kermack, McCrea and Whittaker (1933, §§ 7 and 11).

According to the definition given by those authors, P is at rest relative to S if the direction of the world lines of P and S are parallel with respect to displacement along the null geodesic SP . By (1.8) this implies that $\lambda'' = \lambda'$, and hence $D = 1$.

(c) *Whittaker's first definition of distance*

The first definition of distance, which opened the whole discussion (Whittaker 1931) is more complex than the definition in terms of projected distance, but can be discussed quite simply by the use of optical co-ordinates, taking the base curve to be the world line of the star.

We require two preliminary results, viz. (i) the equation of any space-like surface S_3 can be written in the form $y^4 = \phi(y^1, y^2, y^3)$, and (ii) the area of the cross-section of a pencil of null geodesics starting from the same point A by any S_3 is independent of the orientation of the S_3 .

To establish the first result let $f(y^1, y^2, y^3, y^4) = 0$ be the equation of any surface S_3 . Let $f_\mu = \partial f / \partial y^\mu$ and $f^\mu = h^{\mu\nu} f_\nu$. Then, by (3.11),

$$f^{n-1} = h^{n-1, \mu} f_\mu = f_n.$$

Hence, if $f_n = 0$ at any point P of S_3 , f^{n-1} also vanishes at P . Therefore, for a displacement, $dy^\mu = f^\mu d\sigma$, along the normal to S_3 at P ,

$$ds^2 = h_{pq} f^p f^q (d\sigma)^2.$$

But, by (5.1), this quadratic form is never positive, so that the normal to S_3 at P is space-like or null. Now, if S_3 is a space-like surface, its normal is everywhere time-like. Hence, for a space-like surface, f_n cannot vanish at any point. The equation $f = \text{constant}$ can therefore be solved for y^4 in the form

$$y^4 = \phi(y^1, y^2, y^3).$$

To prove the second lemma, we note that on the null cone of A the metric is given by

$$ds^2 = h_{pq} dy^p dy^q.$$

This null cone intersects any space-like S_3 in a two-way S_2 , whose metric is obtained from that of the null cone by writing

$$dy^4 = \partial\phi/\partial y^i \cdot dy^i.$$

But since the metric on the null cone contains no terms in $dy^n dy^\mu$, this substitution leaves the metric unchanged.

This result, which we have proved in the large, is the generalization of a theorem proved by Walker (1933*a*, § 2) for a thin pencil of null geodesics, given in our notation by

$$a^p \leq y^p \leq a^p + da^p, \quad y^3 = \tau.$$

Our result shows that the area of the cross-section of the pencil by any space-like S_3 is

$$(\det h_{pq})^{\frac{1}{2}} da^1 da^2 = (-h)^{\frac{1}{2}} da^1 da^2, \quad \text{by (3.10).}$$

This area is clearly independent of the orientation of the S_3 and depends only upon the position of the point P near which the pencil meets S_3 . We shall denote this area by V_P , so that

$$V_P = (-h_P)^{\frac{1}{2}} da^1 da^2 = \lambda_P^2 (\eta_P)^{\frac{1}{2}} da^1 da^2, \quad \text{by (3.10).}$$

To obtain Whittaker's original definition of distance, which we shall denote by Δ_W , we consider a second observer Q near the star S , and *at rest relative to the first observer P* . By (6.1) and (6.3) the distance of S from Q in the space of Q is

$$\delta = \lambda' = \lambda''/D = \lambda_Q/D,$$

λ_Q being the y^4 co-ordinate of Q . Then, Whittaker's definition is, in effect,

$$\Delta_W : \delta = V_P^{\frac{1}{2}} : V_Q^{\frac{1}{2}} = (h_P/h_Q)^{\frac{1}{2}}.$$

Therefore

$$\Delta_W = \lambda_P (\eta_P/\eta_Q)^{\frac{1}{2}}/D, \quad (6.4)$$

where Q can be taken at the star S .

(d) *Etherington's distance*

In the modification of Whittaker's original definition due to Etherington (1933) the second observer Q is taken to be near the star S and *at rest relative to the star*. Hence the distance of Q from S is now

$$\delta = \lambda'' = \lambda_Q.$$

Etherington's definition may then be written

$$\Delta_E : \delta = V_P^\dagger : V_Q^\dagger = (h_P/h_Q)^\dagger,$$

whence

$$\Delta_E = \lambda_P(\eta_P/\eta_Q)^\dagger. \quad (6.5)$$

Etherington has proved that Δ_E , unlike Δ_W , is an absolute scalar invariant, which can be expressed in the form

$$\Delta_E = -(g_P g_S / J^2)^\dagger (\partial \Omega / \partial \tau),$$

where g_P, g_S are the values of $\det g_{\mu\nu}$ at P and S in any arbitrary system of co-ordinates, Ω is the characteristic function for S and P ,

$$J = \det (\partial^2 \Omega / \partial x_P^\mu \partial x_S^\nu),$$

and $d\tau$ is an element of the world line of S . (Etherington's proof, which employs Riemannian normal co-ordinates, really presupposes the definition of these co-ordinates for an indefinite manifold, given above in § 2.)

It is obvious that, in general, Δ_W and Δ_E will differ from the projected distances λ_P/D and λ_P , and that, in optical co-ordinates, the variation of η determines the derivation between these different measurements of distance.

(e) *Distance by apparent size*

Etherington has also pointed out that the distance of a star as estimated from its apparent size, say Δ_S , can be obtained from Δ_E by interchanging the roles of the star and the observer. Hence

$$\Delta_S = + (g_S g_P / J^2)^\dagger (\partial \Omega / \partial s_P),$$

where ds_P is an element of the world line of P . Now, by (6.3),

$$D = ds_P / d\tau = - \partial \Omega / \partial \tau : \partial \Omega / \partial s_P.$$

Therefore

$$\Delta_S = \Delta_E / D = \Delta_W, \quad (6.6)$$

so that Whittaker's original definition agrees with the definition in terms of apparent size.

(f) *Luminosity distance*

For the sake of completeness we also add the result obtained by Walker (1933 *b*) that the distance of a star as estimated by its luminosity is given by

$$\begin{aligned} L &= (V_P : V_Q)^{\frac{1}{2}} \delta D \\ &= \lambda_P (\eta_P / \eta_Q)^{\frac{1}{2}} D \\ &= D^2 \Delta_S. \end{aligned} \quad (6.7)$$

(g) *Distance by parallax*

Finally we shall obtain an expression for the distance as determined by observations of parallax, i.e. the angle between the apparent directions of the star as observed by two observers P and \bar{P} .

If the direction of the null geodesic from the star at P is specified by the vector g^μ and if the unit vector along the world line of P is n^μ , then the apparent direction of the star in the instantaneous space of P is specified by the projection of g^μ in this space, i.e. by the vector

$$p^\mu = g^\mu - n^\mu (n_\rho g^\rho).$$

The magnitude of this vector is

$$(p^\mu p_\mu) = -(n_\mu g^\mu)^2.$$

If g^μ is taken to be $dy^\mu/d\lambda$, where λ is the projected distance along the null geodesic, the projection vector being along the world line of the star, then

$$n_\mu g^\mu = n_\mu dy^\mu : d\lambda = \lambda' : \lambda''$$

in the notation of § 6 (*a*). Hence, by (6.3)

$$n_\mu g^\mu = D^{-1}, \quad \text{and} \quad (p^\mu p_\mu) = -D^{-2} \quad (6.8)$$

where D is the Doppler shift at P .

Similarly, if \bar{g}^μ specifies the direction of the null geodesic at \bar{P} , and \bar{n}^μ the unit tangent to the world line of \bar{P} , then the apparent direction of the star at \bar{P} is specified by

$$\bar{p}^\mu = \bar{g}^\mu - \bar{n}^\mu (\bar{n}_\rho \bar{g}^\rho).$$

The general discussion of distance by parallax is somewhat complex as it involves the relative motions of the two observers and the star. We shall therefore restrict ourselves to the simplest case when \bar{P} is at rest in the instantaneous space of P , while the line $P\bar{P}$ in this space is perpendicular

to the apparent direction of the star. If (y^μ) and $(y^\mu + dy^\mu)$ are the co-ordinates of P and \bar{P} respectively these conditions imply that

$$\left. \begin{aligned} \bar{n}^\mu &= n^\mu - \Gamma_{\alpha\beta}^\mu n^\alpha dy^\beta, \\ n_\mu dy^\mu &= 0 \quad \text{and} \quad p_\mu dy^\mu = 0. \end{aligned} \right\} \quad (6.9)$$

The distance by parallax is then given by

$$\Delta_P = dl/d\theta,$$

where dl is the distance of \bar{P} from P , and $d\theta$ is the angle between the apparent directions of the star at P and \bar{P} . It is clear that

$$dl = (-h_{\mu\nu} dy^\mu dy^\nu)^{\frac{1}{2}},$$

but the determination of $d\theta$ is more complicated.

The angle between p^μ at P and \bar{p}^μ at \bar{P} is equal to the angle between $p^\mu + dp^\mu$ at \bar{P} and \bar{p}^μ at \bar{P} , where $p^\mu + dp^\mu$ at \bar{P} is parallel to p^μ at P , i.e.

$$p^\mu + dp^\mu = p^\mu - \Gamma_{\alpha\beta}^\mu p^\alpha dy^\beta.$$

We shall write

$$\begin{aligned} \pi^\mu &= p^\mu + dp^\mu - \bar{p}^\mu \\ &= g^\mu - n^\mu (n_\alpha g^\alpha) - \Gamma_{\alpha\beta}^\mu p^\alpha dy^\beta - \bar{g}^\mu + \bar{n}^\mu (\bar{n}_\alpha \bar{g}^\alpha). \end{aligned}$$

Now, in optical co-ordinates, the components of g^μ and \bar{g}^μ are each $(0, 0, 0, 1)$. Hence, by (6.9)

$$\begin{aligned} \pi^\mu &= -\Gamma_{\alpha\beta}^\mu g^\alpha dy^\beta - \Gamma_{\alpha\beta, 4}^\mu n^\alpha n^\beta dy^\beta \\ &= \frac{1}{2} \frac{\partial h_{\alpha\beta}}{\partial \lambda} (n^\mu n^\alpha - h^{\mu\alpha}) dy^\beta, \end{aligned} \quad (6.10)$$

neglecting products of the displacements dy^α .

It remains to calculate the angle $d\theta$ between \bar{p}^μ and $\bar{p}^\mu + \pi^\mu$. The standard formula for the angle between two vectors E^μ and F^μ gives the cosine of the angle as

$$\cos \theta = (E_\mu F^\mu) (E_\mu E^\mu)^{-\frac{1}{2}} (F_\mu F^\mu)^{-\frac{1}{2}},$$

but for small angles it will be better to form an expression for the sine of the angle. Writing

$$R^{\mu\nu} = E^\mu F^\nu - E^\nu F^\mu,$$

we easily find that

$$\sin^2 \theta = \frac{1}{2} (R^{\mu\nu} R_{\mu\nu}) (E_\mu E^\mu)^{-1} (F_\mu F^\mu)^{-1}.$$

Now, if

$$E^\mu = \bar{p}^\mu, \quad \text{and} \quad F^\mu = \bar{p}^\mu + \pi^\mu,$$

then

$$R^{\mu\nu} = \bar{p}^\mu \pi^\nu - \bar{p}^\nu \pi^\mu,$$

$$\text{and} \quad \frac{1}{2} R^{\mu\nu} R_{\mu\nu} = (\bar{p}^\mu \bar{p}_\mu) (\pi^\nu \pi_\nu) - (\bar{p}^\mu \pi_\mu)^2.$$

$$\begin{aligned} \text{But, by (6.10),} \quad \bar{p}_\mu \pi^\mu &= -\frac{1}{2} \dot{h}_{\alpha\beta} \bar{p}^\alpha dy^\beta \\ &= +\frac{1}{2} \dot{h}_{\alpha\beta} \bar{n}^\alpha dy^\beta (\bar{n}_\mu \bar{g}^\mu), \end{aligned}$$

using dots to denote differentiation with respect to λ .

$$\text{Also} \quad \pi_\mu \pi^\mu = \frac{1}{4} \dot{h}_{\alpha\beta} \dot{h}_{\rho\sigma} (h^{\alpha\rho} - n^\alpha n^\rho) dy^\beta dy^\sigma.$$

$$\text{Hence} \quad \frac{1}{2} R^{\mu\nu} R_{\mu\nu} = -\frac{1}{4} (\bar{n}_\mu \bar{g}^\mu)^2 \dot{h}_{\alpha\beta} \dot{h}_{\rho\sigma} h^{\alpha\rho} dy^\beta dy^\sigma.$$

$$\text{Now, by (6.8),} \quad E^\mu E_\mu = \bar{p}^\mu p_\mu = -\bar{D}^2,$$

where \bar{D} is the Doppler shift at \bar{P} ; and

$$\begin{aligned} F^\mu F_\mu &= \bar{h}_{\mu\nu} (p^\mu + dp^\mu) (p^\nu + dp^\nu) \\ &= h_{\mu\nu} p^\mu p^\nu, \quad (\text{since the displacement is "parallel"}), \\ &= -D^2. \end{aligned}$$

Therefore, inserting the values of $\frac{1}{2} R^{\mu\nu} R_{\mu\nu}$, $E_\mu E^\mu$ and $F_\mu F^\mu$ in the expression for $\sin^2 \theta$, we find that

$$d\theta = \frac{1}{2} D \{ -\dot{h}_{\alpha\beta} \dot{h}_{\rho\sigma} h^{\alpha\rho} dy^\beta dy^\sigma \}^{\frac{1}{2}}.$$

The distance by parallax is therefore

$$\Delta_P = \frac{2(-h_{\mu\nu} dy^\mu dy^\nu)^{\frac{1}{2}}}{D(-\dot{h}_{\alpha\beta} \dot{h}_{\rho\sigma} h^{\alpha\rho} dy^\beta dy^\sigma)^{\frac{1}{2}}}. \quad (6.11)$$

7. OPTICAL CO-ORDINATES IN AN EXPANDING UNIVERSE

A standard form for the metric of an expanding universe is

$$ds^2 = dt^2 - R^2(t) (1 + \frac{1}{4} k r^2)^{-2} (dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2),$$

which is conveniently written as

$$ds^2 = dt^2 - R^2(t) \{ du^2 - U^2(u) (d\theta^2 + \sin^2 \theta d\phi^2) \},$$

$$\text{where} \quad u = \int_0^r (1 + \frac{1}{4} k r^2)^{-1} dr \quad \text{and} \quad U(u) = r(1 + \frac{1}{4} k r^2)^{-1}. \quad (7.1)$$

We shall construct a system of optical co-ordinates valid for all values of k and $R(t)$.

We take the basic curve C to be the time-like geodesic $r = 0$ or $u = 0$. The null geodesics through any arbitrary point A ($t = \tau$) on C are then given by

$$\theta = \text{constant}, \quad \phi = \text{constant},$$

$$u = f(t) - f(\tau), \quad \text{where} \quad f(t) = \int^t \frac{dx}{R(x)}. \quad (7.2)$$

It is clear that θ , ϕ and τ (as defined by the last equation) are optical co-ordinates corresponding to y^1 , y^2 , and y^3 of § 3.

To obtain the remaining optical co-ordinate λ we have to determine the projected length of the null geodesic (7.2). In an obvious notation we shall write the components of the projection vector as A^t , A^u , A^θ , A^ϕ and the three-index symbols as Γ_{tu}^θ etc. It is clear that $A^\theta = 0$ and $A^\phi = 0$. The remaining components are determined by the equations

$$dA^t = -\Gamma_{uu}^t A^u du = -RR' A^u du,$$

$$dA^u = -\Gamma_{ut}^u (A^u dt + A^t du) = -R^{-1}R' (A^u dt + A^t du),$$

and the initial conditions,

$$A^t = 1, \quad A^u = 0, \quad \text{when} \quad u = 0.$$

We simplify these equations by writing $dt = R du$ (since the curve is a null geodesic) and by taking $R(t)$ as a new independent variable. We thus obtain

$$\frac{dA^t}{dR} = -A^u, \quad \frac{dA^u}{dR} = -\frac{A^u}{R} - \frac{A^t}{R^2}.$$

Hence

$$A^t = \frac{1}{2} \left\{ \frac{R(t)}{R(\tau)} + \frac{R(\tau)}{R(t)} \right\},$$

$$A^u = \frac{1}{2} \left\{ \frac{-1}{R(\tau)} + \frac{R(\tau)}{[R(t)]^2} \right\}.$$

This solution is clearly valid even if $R(t)$ is a mere constant.

The projected length of the null geodesic is now given by

$$\lambda = \int_{u=0} \{ A^t dt - R^2(t) A^u du \},$$

the integral being taken along the curve (7.2), so that $dt = R du$. Hence

$$\lambda = \frac{g(t) - g(\tau)}{R(\tau)}, \quad \text{where} \quad g(t) = \int^t R(x) dx. \quad (7.3)$$

We now have, from (7.2) and (7.3),

$$d\tau = \frac{R(\tau)}{R(t)} \{dt - R(t) du\}$$

and

$$d\lambda = \frac{R(t) dt - R(\tau) d\tau}{R(\tau)} - \frac{\lambda R'(\tau)}{R^2(\tau)} d\tau.$$

Hence, if

$$T = 2 - \frac{R^2(t)}{R(\tau)} + \frac{2\lambda R'(\tau)}{R^2(\tau)}, \quad (7.4)$$

then

$$T d\tau + 2d\lambda = \frac{R(t)}{R(\tau)} \{dt + R(t) du\}.$$

Therefore

$$T d\tau^2 + 2d\tau d\lambda = dt^2 - R^2(t) du^2.$$

Hence, in optical co-ordinates,

$$ds^2 = T d\tau^2 + 2d\tau d\lambda - R^2(t) U^2(u) (d\theta^2 + \sin^2 \theta d\phi^2). \quad (7.5)$$

The simplicity of this expression is apparently spoilt by the complicated form of T , but fortunately T does not appear in the expression for the determinant of the coefficients, $h = \det h_{\mu\nu}$, whose value is

$$h = -R^4(t) U^4(u) \sin^2 \theta.$$

Near the base curve C , r and u are small, and we have the following approximations:

$$u \doteq \frac{t - \tau}{R(\tau)}, \quad \text{from (7.2),}$$

$$\lambda \doteq t - \tau, \quad \text{from (7.3),}$$

$$U(u) \doteq r \doteq u, \quad \text{from (7.1),}$$

$$h \doteq -\lambda^4 \sin^2 \theta.$$

Therefore

$$\eta_A = \lim_{\lambda \rightarrow 0} (-h/\lambda^4) = \sin^2 \theta,$$

and

$$(-h_P/\eta_A)^{\frac{1}{2}} = U(u_P) R(t_P) = \frac{R(t_P) \cdot r_P}{1 + \frac{1}{4} k r_P^2}. \quad (7.6)$$

We can now write down explicit expressions for the various astronomical distance functions from a star whose world line is $r=0$ to an observer at (t, r, θ, ϕ) . The projected distance is

$$\lambda = \int_{\tau}^t \frac{R(x) dx}{R(\tau)}.*$$

* This does not agree with the expression given by McCrea (1935, eq. (11)), which appears to be erroneous.

where τ is defined by the equation

$$\int_0^r \frac{dx}{1 + \frac{1}{4}kr^2} = \int_\tau^t \frac{dx}{R(x)}.$$

If the observer is at rest in the (t, r, θ, ϕ) co-ordinate system, the Doppler shift is

$$\begin{aligned} D &= ds/d\tau && \text{by (6.1)} \\ &= dt/d\tau \\ &= R(t)/R(\tau)^* && \text{by (7.2).} \end{aligned}$$

Etherington's distance is

$$\Delta_E = (-h_P/\eta_A)^{\frac{1}{2}} = \frac{r R(t)}{1 + \frac{1}{4}kr^2}, \quad \text{by (7.6)}$$

Whittaker's first distance function, or the distance by apparent size, is

$$\Delta_E/D = \frac{r R(\tau)}{1 + \frac{1}{4}kr^2} \cdot \dagger$$

The luminosity distance is

$$\Delta_E D = \frac{R^2(t)}{R(\tau)} \frac{r}{1 + \frac{1}{4}kr^2} \cdot \ddagger$$

Finally we shall calculate the distance by parallax in the simplest case when P and \bar{P} are at rest in the original co-ordinate system and make their observations simultaneously in this system. The original co-ordinates of P and \bar{P} will then be (t, r, θ, ϕ) and $(t, r, \theta + d\theta, \phi + d\phi)$. The distance by parallax then becomes, by (6.11),

$$\begin{aligned} \Delta_P &= \frac{2(-h_{pq}dy^p dy^q)^{\frac{1}{2}}}{D(-\dot{h}_{pq}\dot{h}_{rs}\dot{h}^{pr}dy^q dy^s)^{\frac{1}{2}}} \\ &= f \div (D \partial f / \partial \lambda), \end{aligned}$$

where

$$f = R(t) U(u).$$

Now, after a little calculation, we find that

$$\frac{\partial f}{f \partial \lambda} = \frac{R(\tau)}{[R(t)]^2} \left\{ R'(t) + \frac{U'(u)}{U(u)} \right\},$$

whence

$$\Delta_P = \frac{r R(t)}{1 - \frac{1}{4}kr^2 + r R'(t)} \cdot \S$$

* Cf. Tolman (1934, eqn. (178.5)) and Walker (1933b, eqn. (4.1)), where the observer is at the origin.

† Cf. Walker (1933b, eqns. (37) and (40)), where R corresponds to our $R(\tau)$.

‡ Cf. McCrea (1935, eqn. (15)), where R_0 corresponds to our $R(t)$ and R_1 to our $R(\tau)$.

§ Cf. McCrea (1935, eqn. (10)), where R_0 corresponds to our $R(t)$.

SUMMARY

This paper develops the theory of a new system of normal co-ordinates (y^μ) in an n -dimensional manifold with an indefinite metric of signature $+1-(n-1)$. In this system of co-ordinates the expression for the metric has the form

$$ds^2 = h_{jk} dy^j dy^k + 2dy^{n-1} dy^n, \quad (j, k = 1, 2, \dots, n-1).$$

Any arbitrary time-like curve C being taken as the base curve, the forward null cone at any point A on C is given by $y^{n-1} = \text{constant}$, C itself is given by the single equation $y^n = 0$, and any forward null geodesic G from A is given by $y^j = \text{constant}$ ($j = 1, 2, \dots, n-1$). y^n is the "projected length" defined in a general co-ordinate system (x^μ) by

$$y^n = \int K_\mu dx^\mu,$$

where the integral is taken along G from A , and K_μ is a unit vector transported along G by parallel displacement and forming the tangent to C at A .

Hadamard's theory of null geodesics as bicharacteristics leads to a generalization of this system of normal co-ordinates in which $y^{n-1} = \text{constant}$ is the equation of a system of wave fronts, $y^j = \text{constant}$ are the equations of the orthogonal null geodesics or rays, and y^n is a projected length measured along the rays.

The theory is applied to optical problems in general relativity. Fermat's principle of minimum time is established rigorously for any relativistic manifold. The discussion of astronomical determinations of distance is simplified by the use of the new system of normal co-ordinates, which gives explicit expressions for the various distance functions. These general results are illustrated by the case of a manifold representing an isotropic expanding universe. Incidentally the use of normal co-ordinates of Riemann's type in relativistic optics is justified by a new definition in terms of "projected distances".

REFERENCES

- Etherington, I. M. H. 1933 *Phil. Mag.* (7), 15, 761.
 Hadamard, J. 1903 "Leçons sur la propagation des ondes."
 — 1923 "Lectures on Cauchy's Problem..."
 Kermack, W. O., McCrea, W. H. and Whittaker, E. T. 1933 *Proc. Roy. Soc. Edinb.* 53, 31.
 McCrea, W. H. 1935 *Z. Astrophys.* 9, 290.

- Ruse, H. S. 1929 *Proc. Lond. Math. Soc.* **32**, 87.
— 1932 *Proc. Roy. Soc. Edinb.* **52**, 183.
— 1933 *Proc. Roy. Soc. Edinb.* **53**, 79.
Tolman, R. C. 1934 "Relativity, Thermodynamics and Cosmology."
Walker, A. G. 1933*a* *Quart. J. Math.* **4**, 71.
— 1933*b* *Mon. Not. R. Astr. Soc.* **94**, 159.
Whittaker, E. T. 1931 *Proc. Roy. Soc. A*, **133**, 93.
-

The separation of isotopes for the investigation of nuclear transmutations

BY E. LEIGHTON YATES, PH.D.

(Communicated by E. V. Appleton, F.R.S.—Received 27 June 1938)

INTRODUCTION

The bombardment of an element with intense beams of protons or deuterons may lead to a variety of processes of nuclear transmutation. When, as is normally the case, the bombarded element is composed of two or more isotopes it is often difficult to establish with certainty the nature of the transmutations which give rise to the observed disintegration products. The use of separated isotopes as targets for such experiments greatly facilitates the interpretation of the results. The investigation of the radioactivity of ${}^6\text{Li}$, for example, was simplified considerably by the use of the separated isotopes of lithium (Rumbaugh and Hafstad 1936; Lewis, Burcham and Chang 1937). Apart from hydrogen, lithium is the only light element which has been separated into its constituent isotopes in quantities sufficient for disintegration experiments. Oliphant, Shire and Crowther (1934) have described two simple mass spectrographs by means of which quantities of the lithium isotopes of the order of 10^{-7} g. were obtained, while Rumbaugh (1936) has used a mass spectrograph of more complicated design to produce several hundred micrograms of the lithium and potassium isotopes. The present paper describes an improved but still comparatively simple mass spectrograph, similar to those described by Oliphant, Shire and Crowther, which has been used for the separation of $30\mu\text{g}$. of the heavier isotope of lithium and $2\mu\text{g}$. of each of the isotopes of boron. The apparatus has also been used for the preparation of several micrograms of the isotope of carbon of mass 12 free from the heavier isotope.

THE ION SOURCES

(a) *Lithium*

Lithium ions may be obtained by heating a platinum filament coated with a lithium salt. The preparation of the isotopes demands, however, a stable ion source capable of yielding a copious supply of lithium ions over long periods. Blewett and Jones (1936) have coated a platinum filament 1 cm.² in area with β -eucryptite, $\text{Li}_2\text{O} : \text{Al}_2\text{O}_3 : 2\text{SiO}_2$, and obtained lithium ion currents of the order of 1 mA. In some experiments in the Cavendish Laboratory, Shire had designed a platinum filament which held larger quantities of the ion source than the simple coated strips described in the

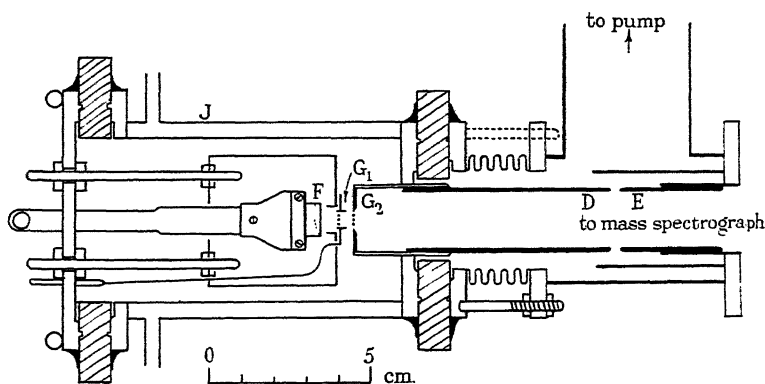


FIG. 1

published work (Oliphant, Shire and Crowther 1934). This design also enabled the source to be maintained at a temperature more nearly equal to that of the platinum filament. The front face of a platinum V-shaped trough, 2 cm. long and 1 cm. deep, was pierced with small holes through which the semi-fluid mixture diffused to provide the emitting surface. The temperature was maintained by a current of about 65 amp. through the trough and, since the emission increased rapidly with the temperature, the latter was kept near to the melting-point of platinum. Owing, therefore, to the risk of breakdown, molybdenum and tungsten troughs were tried, but in each case the metal was attacked by the mixture at the higher temperatures. At lower temperatures the yield was definitely less than that obtained with platinum. The initial separation of the ions was obtained by applying 250 V between the front face of the platinum trough, *F* (fig. 1), and a grid, *G*₁, of fine tungsten wire placed 5 mm. from it. The second stage consisted of about 2000 V applied between the grid *G*₁ and a similar grid,

G_2 , mounted in the end of a steel cylinder which formed part of the ion lens, $D-E$. This lens was used to focus the ions on the exit slit of the mass spectrograph, about 30 cm. distant. The potential across the ion lens gap, $D-E$, was about 6000 V, the focus being adjusted by varying the potential on G_2 .

The ion source fitted inside a steel water-jacket, J , and the accelerating grids and ion lens were completely enclosed in steel tubes. These steel tubes shielded the ion beam from the stray field of the magnet which was employed in the mass spectrograph. A short length of sylphon bellows allowed the ion source to be moved relative to the mass spectrograph against which it fitted. In this way the maximum ion beam was directed into the mass spectrograph. The apparatus was evacuated by means of a large mercury diffusion pump.

The maximum ion current obtainable from the source was about 2 mA. After focusing, however, the total ion current was never more than one-tenth of this, giving about 40 μ A of the ions of the heavier isotope of lithium. With this current the isotope was deposited at the rate of 10 μ g./hr. Under these conditions the source did not last for more than three hours and deposits of more than 30 μ g. were not obtained in a single run. After removal from the apparatus, the deposit of lithium isotope was fixed by exposure to the vapour of hydrochloric acid.

(b) *Boron*

The main difficulty in the separation of the boron isotopes is the provision of a suitable ion source. Several attempts were made to obtain boron ions by heating boron salts both alone and in combination with other salts, but in no case was a measurable boron isotope current obtained. An attempt was also made to produce the ions in a high frequency spark, as used by Dempster (1936) and his collaborators (Sheng-Lin Ch'u 1936), but without success. Aston (1920) has obtained boron ions by introducing boron trifluoride into a gas discharge. The yield given by the simple discharge tube used by Aston would not be sufficient to justify its use in the preparation of samples of the isotopes. The simple gas discharge produces ions of variable energy and the addition of a subsidiary electrostatic field to produce the homogeneity in energy required for the successful operation of the velocity selector would make the apparatus unmanageably complicated and would certainly result in the loss of a large proportion of the ions. These difficulties were avoided by using a low voltage arc similar to that used by Oliphant (1929) for the production of helium ions. It was found that the addition of boron trifluoride to the helium resulted in the production of boron ions.

Two types of low voltage arc were employed. In the preliminary experiments an open arc was used but gave a low yield of boron ions. Modification of this arc to one of the capillary type increased the boron ion current by a factor of ten. The arc was similar to that used by Tuve, Dahl and Hafstad (1935) for the production of hydrogen ions. The use of boron trifluoride in the arc makes it necessary to use tungsten instead of coated filaments, and the design of Tuve, Dahl and Hafstad was therefore modified to allow of efficient water-cooling round the arc. The essential details are shown in fig. 2. The steel capillary, *C*, was made 2 cm. long and 3 mm. in diameter,

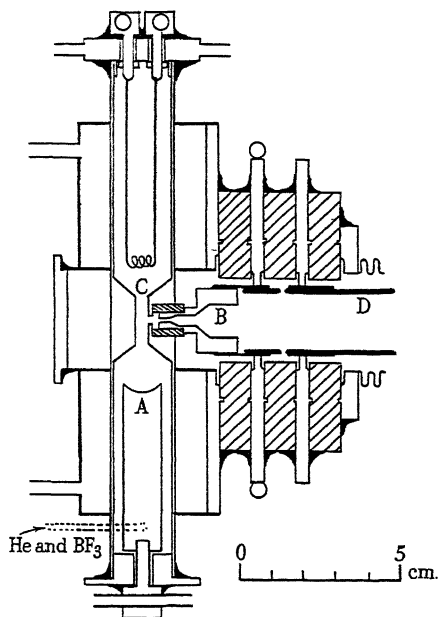


FIG. 2

but all attempts to strike the arc in helium and in hydrogen failed. Increasing the diameter of the capillary to 4 mm. had no effect, but when the capillary was shortened to 15 mm. no difficulty was experienced in striking the arc to the anode, *A*. The arc was first struck in pure helium and the boron trifluoride slowly added. The helium was then gradually reduced until, after about 1 hr., the arc would run in boron trifluoride alone. In order to freeze out mercury and hydrocarbon vapours, the gases supplied to the arc were passed through a trap immersed in solid carbon dioxide and acetone. The rate of flow of the helium was controlled with a needle valve. The use of such a valve to control the flow of the boron trifluoride was not possible

owing to the chemical activity of the gas, and it was therefore allowed to leak through fine glass capillaries from a small reservoir containing the gas at a pressure of about 10 cm. The gas in the reservoir was stored over mercury.

The largest ion currents were obtained with an arc current of 0.6 amp. and a potential difference across the arc of about 60 V. A subsidiary arc running between the filament and the metal capillary helped to stabilize the main arc; this subsidiary arc ran at 0.1 amp. and maintained the capillary at a positive potential of 20 V with respect to the filament. The ions were drawn out of the arc through a 2 mm. hole in the centre of the capillary, by a negative potential of 3000 V on the probe, *B*. This probe was held in position by means of a water-cooled steel disk and was insulated from the ion source by a cylinder of fired steatite. The hole through the centre of the probe was 4 mm. long and 1 mm. in diameter. The current to the probe was about 0.5 mA and the ion current through the probe was about 10 % of this. After emerging from the probe canal the ions were focused by two ion lenses, the first between the probe and the cylinder *D*, and the second between the two cylinders *D* and *E* (fig. 1). The potential of the cylinder *D*, was found to need critical adjustment in order to obtain an image of the probe canal at the exit slit of the mass spectrograph. The most satisfactory potentials on the cylinders *E* and *D*, on the probe *B*, and on the capillary *C*, were found to be 0, 8500, 7000 and 10,000 V respectively. Except for the insulation of glass and steatite and for the brass water-cooling tubes, the apparatus was constructed of steel thus ensuring the protection of the ion beam from the stray field of the magnet. The ion lens was evacuated with a four stage mercury diffusion pump.

The ion currents obtained with this arc were considerably greater than those obtained with the open arc. A current of $2\text{ }\mu\text{A}$ of the isotope of mass 11 was obtained producing deposits of about $2\text{ }\mu\text{g.}$ in $2\frac{1}{2}$ hr. The ion current of the lighter isotope was one-quarter of this and deposits of $1\text{ }\mu\text{g.}$ were obtained by collecting for 5 hr.

THE MASS SPECTROGRAPH

The mass spectrograph made use of the device commonly known as a velocity selector in which ions of a definite velocity are separated by the simultaneous action of crossed electrostatic and magnetic fields. This type of mass spectrograph has advantages over most other types when used to prepare samples of separated isotopes. The ions of the required isotope travel along approximately straight lines and the apparatus is therefore easy to

construct and to adjust. The separation of two isotopes is proportional to the magnetic field and the resolution is thus under control. The separation of two beams of ions, of masses m_1 and m_2 and charge e (e.s.u.), after traversing L cm. in a region of horizontal magnetic field of intensity H gauss, and vertical electrostatic field, is given by the relation,

$$d = \frac{L^2 H}{2c} \left(\frac{1}{\sqrt{m_1}} - \frac{1}{\sqrt{m_2}} \right) \sqrt{\left(\frac{e}{2V} \right)},$$

where c is the velocity of light and V (e.s.u.) is the potential through which the ions have passed before entering the selector. Two factors limited the choice of this potential, V . The production of the ions demanded a certain minimum value, and also an increase in the energy of the ions reduced the effect of space charge on the width of the beam. The accelerating potential therefore, was determined by these considerations. The magnetic field was then chosen to ensure that the linear separation of the isotopes was several times greater than the width of the slit through which the ions emerged from the velocity selector.

The main features of the mass spectrograph were similar to those described by Oliphant, Shire and Crowther (1934). There were two improvements in the design which simplified the working of the apparatus. The flat parallel plates P (fig. 3), between which the electrostatic field was maintained, were mounted as a separate unit which could be withdrawn from the spectrograph for cleaning and adjustment. Considerable improvement also was made in the method of collection of the isotopes. On emerging from the spectrograph the ions entered a copper box, B . In the lithium separation experiments this box was suspended from a glass tube containing liquid air, the deposits being obtained upon a copper surface, A , cooled to liquid air temperatures. When the collecting foil was cooled with liquid air in the presence of boron trifluoride the gas solidified on the surface and contaminated the isotopes deposited upon it. The box was therefore supported on glass pillars and was not cooled.

The collecting box was raised to a potential of about 1000 V less than the total potential employed in the acceleration of the ions. This reduced the energy of the ions before deposition, thus minimizing the effect of sputtering and reducing the secondary electron emission. The foil upon which the isotopes were deposited was supported on a removable tray and the rate of deposition of the isotope was examined during an experiment by measuring the current flowing between the foil and the box. Owing to the retarding potential of the collecting box and also to the presence of the stray field of the magnet, the ions were deviated out of the neutral beam. A copper

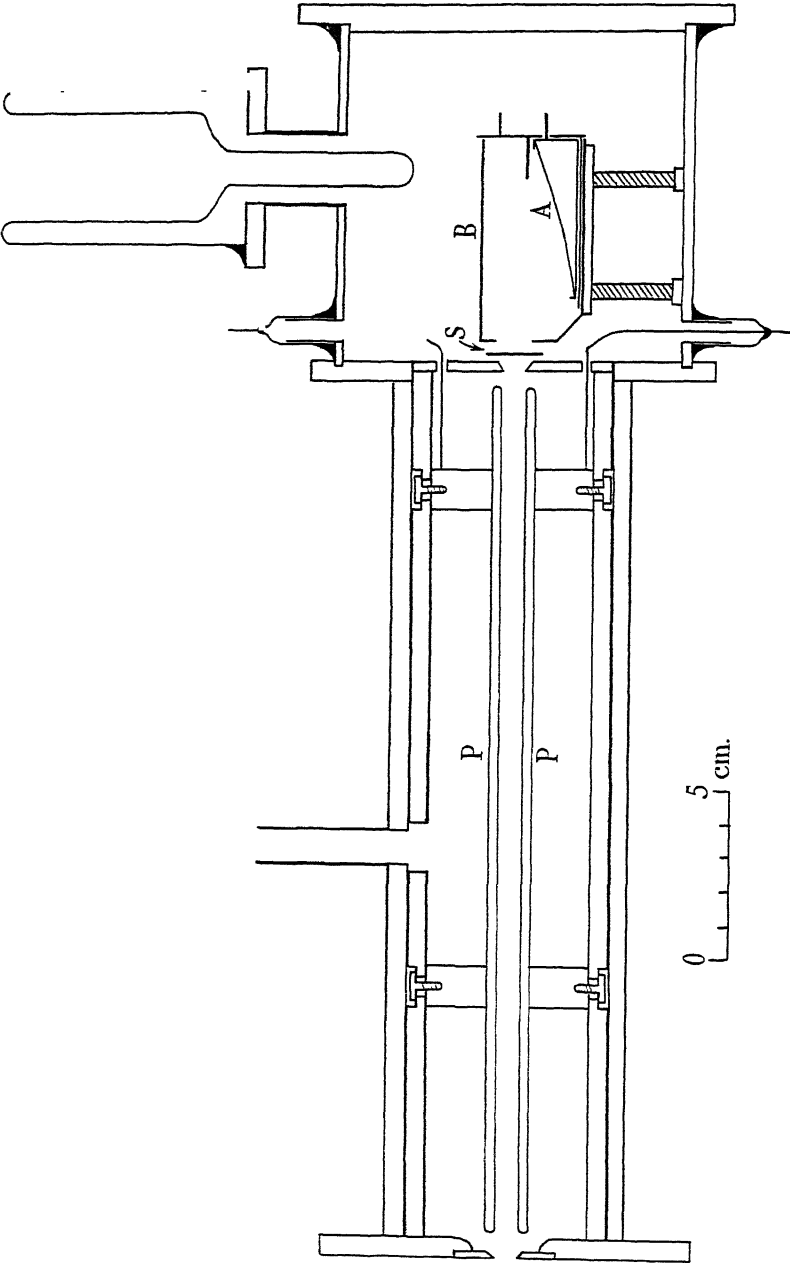


FIG. 3

shutter, *S*, was arranged to close the exit slit of the mass spectrograph by the rotation of a ground seal in the end plate of the apparatus. The required isotope was collected on the shutter to avoid contamination of the collecting foil while the initial adjustments were being made. When these adjustments had been completed, the shutter was opened and the isotope allowed to fall directly on to the collecting foil. A small mercury diffusion pump was connected to the apparatus adjoining the collecting box to reduce the contamination of the collecting surface by the boron trifluoride gas.

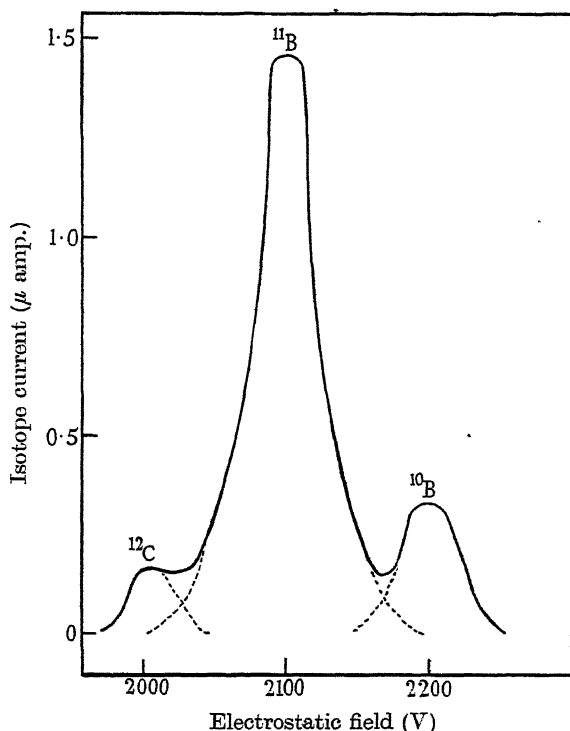


FIG. 4

The magnet pole pieces, which were 25 cm. long and 6 cm. deep, fitted tightly against the sides of the brass box containing the electrostatic field assembly. The width of this box was 3 cm. The copper plates, *P*, 2 cm. wide and 25 cm. long, mounted 7 mm. apart, were made parallel by suitable washers. The potential for the electrostatic field was obtained from a small "voltage-doubler" circuit capable of providing ± 2000 V, the centre point being earthed to the mass spectrograph. The ion beam was limited by slits, 15 by 5 mm., at each end of the mass spectrograph. When the magnetic

field was 5000 gauss and the accelerating potential was 8000 V, the electrostatic fields necessary to collect the isotopes of lithium were 2500 V/cm. for ${}^6\text{Li}$ and 2300 V/cm. for ${}^7\text{Li}$. In the case of boron, where an accelerating potential of 10,000 V and a magnetic field of 7500 gauss were employed, the electrostatic fields were 3150 V/cm. and 3000 V/cm. for ${}^{10}\text{B}$ and ${}^{11}\text{B}$ respectively.

The power of the velocity selector to separate isotopes depends upon the constancy in energy of the ions. A variation of 9 % in the accelerating potential of the ions produces an overlapping of the spectral lines and is sufficient to prevent separation of the boron isotopes. It was found that increasing the potential on the probe (B , fig. 2), resulted in an increased isotope current but reduced the resolution owing to its producing a greater variation in the energy of the ions entering the mass spectrograph. In order that a deposit of the order of one microgram of ${}^{10}\text{B}$ could be obtained in one day it was necessary to use a potential between the capillary and the probe of 3000 V and a total accelerating potential of 10,000 V. The resolution was restored by increasing the magnetic field to 7500 gauss. Under such conditions the variation of boron isotope current with the electrostatic field is shown in fig. 4.

RESULTS

(a) *Lithium*

Several samples of both isotopes have been subjected to artificial disintegration experiments in the High Tension Laboratory (Lewis, Burcham and Chang 1937). Bearing in mind that normal lithium contains 91.7 % of ${}^7\text{Li}$ (Aston 1933), it was found that the samples were of high purity. The disintegration experiments indicated that on the ${}^6\text{Li}$ sample the amount of ${}^7\text{Li}$ was less than 2.5 % of the amount of this isotope normally associated with ${}^6\text{Li}$; while on the ${}^7\text{Li}$ sample the amount of ${}^6\text{Li}$ was less than 1.0 % of the natural amount of this isotope.

(b) *Boron*

The boron isotopes have been tested by bombardment with high energy protons and deuterons in the High Tension Laboratory. A portion of copper foil, which had been in the apparatus adjacent to the target upon which the boron isotope of mass 11 had been deposited, was also bombarded to test for the general deposit of boron previously found when the foil was cooled with liquid air. From these experiments the purity of the targets could be estimated. A deposit of boron was found upon the copper foil amounting to about 7 % of the isotope deposits. The isotopes were therefore deposited

upon surfaces already contaminated by boron. Even had the mass separation been perfect this contamination limited the purity of the isotopic sample. It is difficult to see how this contamination can be avoided, for the instability of boron trifluoride gives rise to a deposit of boron throughout the apparatus. About 2.7 % of the isotope of mass 10 was found on the target upon which the heavier isotope was deposited. This was almost entirely due to the general contamination. The amount of ^{11}B in the ^{10}B deposit was 15 %, about 5 % of this being due to the contamination. The isotope of mass 10 being four times rarer than the heavier isotope, it was to be expected that the purity of the deposit of this isotope would be less than that of ^{11}B . Although the isotope deposits were not of very high purity they were sufficiently pure for the purpose for which they were required. It may be observed that the percentage of ^{10}B in the separated sample was more than four times greater than in natural boron.

The experiments on the transmutations of the separated isotopes will be described in a future communication.

(c) Carbon

The same apparatus has been used for the preparation of samples of the carbon isotope of mass 12. A 50 % mixture of helium and heavy methane (CD_4), was introduced into the capillary arc. It was found that the use of helium increased the carbon isotope current by a factor of two. Heavy methane was used in order to avoid the large current of ions of mass 13 ($^{12}\text{C}^1\text{H}$), which would have resulted from the use of ordinary methane. Ion currents of $3\text{ }\mu\text{A}$ of the desired isotope were readily obtained and deposits of $2\text{ }\mu\text{g}$. were prepared by collecting for about $1\frac{1}{2}$ hr. No attempt has yet been made to collect samples of the carbon isotope of mass thirteen.

I wish to acknowledge my gratitude to the Department of Scientific and Industrial Research for an award which enabled me to carry out this work.

SUMMARY

A mass spectrograph of simple design has been used to prepare deposits containing a few micrograms of the separated isotopes of lithium, boron and carbon. In the case of lithium, a filament type ion source was employed, while the boron and carbon ions were produced in a low voltage capillary arc running in boron trifluoride and in a mixture of heavy methane and helium respectively. On examination in artificial disintegration experiments, the samples were found to be of high purity.

REFERENCES

- Aston 1920 *Phil. Mag.* **40**, 628.
— 1933 "Mass Spectra and Isotopes."
Blewett and Jones 1936 *Phys. Rev.* **50**, 464.
Dempster 1936 *Rev. Sci. Instrum.* **7**, 46.
Lewis, Burcham and Chang 1937 *Nature, Lond.*, **139**, 24.
Oliphant 1929 *Proc. Roy. Soc. A*, **124**, 228.
Oliphant, Shire and Crowther 1934 *Proc. Roy. Soc. A*, **146**, 922.
Rumbaugh 1936 *Phys. Rev.* **49**, 882 (A).
Rumbaugh and Hafstad 1936 *Phys. Rev.* **50**, 681.
Sheng-Lin Ch'u 1936 *Phys. Rev.* **50**, 212.
Tuve, Dahl and Hafstad 1935 *Phys. Rev.* **48**, 241.
-

The scattering of fast β -particles by mercury nuclei

BY A. BARBER AND F. C. CHAMPION

King's College, London

(Communicated by C. D. Ellis, F.R.S.—Received 17 February 1938—

Revised 28 June 1938)

[Plate 2]

1. INTRODUCTION

An account of the scattering of fast β -particles emitted by Ra E, with energies lying in the region of 1 mV by nitrogen nuclei has been given in a previous paper (Champion 1936). Using an expansion chamber it was found that the scattering between 20 and 180° was in fair agreement with a formula deduced by Mott (1932) on wave mechanics. The present paper gives an account of similar experiments carried out with mercury nuclei. Apart from very large deviations from Mott's formula, other phenomena of an interesting nature are observed.

2. EXPERIMENTAL

A standard semi-automatic expansion chamber of rubber piston type was used and a magnetic field of 400 oersteds was supplied by a pair of Helmholtz coils in the usual way. Illumination was provided by two groups of four 100 V 100 W semicircular filament lamps, flashed on a 200 V D.C. circuit for a fraction of a second. The lamps were silvered on the surface behind the filament and covered with black Chinese lacquer, except for a slit some 3 mm. wide in the direction of the chamber. They were run continuously at dull-red heat, the resistance being short-circuited at the instant of flash. Used in this way, each lamp lasted for more than a thousand photographs. Mercury was introduced into the expansion chamber in the form of mercury dimethyl vapour. The mercury dimethyl, which is liquid at room temperature, was contained in a U-tube which communicated through a tap with a hole in the top of the chamber. In order to make sure that the chamber was kept saturated with vapour, opportunity was frequently provided for fresh vapour to be drawn into the chamber, by opening the communicating tap and operating the chamber piston before the

commencement of each run of fifty photographs. The vapour pressure of mercury dimethyl was measured and found to be 3.65 cm. of mercury at 18.5° C. The remaining gas in the chamber was nitrogen, and since the chamber was operated at atmospheric pressure, the gas mixture was about 95 % nitrogen and 5 % mercury dimethyl vapour. The disadvantage of a relatively small number of mercury nuclei for scattering purposes is counterbalanced by their high nuclear charge Z , since the scattering would be expected to be proportional to Z^2 . A higher percentage of mercury nuclei would be a doubtful advantage when using β -particles of 1 mV energy, since the accompanying increase in the number of *small* nuclear deflexions would render the determination of the velocities of the β -particles from the curvatures of the tracks in a magnetic field very uncertain.

The analysis of the photographs was carried out by replacing the developed plates in the cameras and reprojecting the images of the tracks on to white screens capable of orientation in any plane. Angles of deflexion were determined by laying fine steel hat-pins along the reconstructed image of the collision and then using a protractor. Velocities were determined from the curvatures of the tracks, which were measured by fitting circles of known curvature drawn in Indian ink on stiff, transparent cellophane. The method was tested using artificial tracks of curved wire in which were made sharp bends of from 20 to 150°, in different planes. Measurement of the reprojected images gave results accurate to within one or two degrees.

3. RESULTS

Some 2000 pairs of photographs were taken, yielding about 350 m. of track of β -particles with velocities between 0.86 and 0.95 that of light. As in the previous experiments with pure nitrogen (Champion 1936) only those nuclear deflexions were counted in which the deviation was in an opposite direction to the original direction of curvature of the incident β -particle. The collision is then indicated very clearly as a cusp in an otherwise smooth curve. The results for the elastic collisions are shown in Table I.

TABLE I. ELASTIC COLLISIONS

θ	Total observed	N calculated	Difference due to Hg	Mott
20–30°	80	43	37	276
30–60°	58	28	30	180
60–180°	14	6	8	41
20–180°	152	77	75	497

The first column contains arbitrarily chosen angular limits, while the second column gives the observed numbers of β -particles scattered within these limits. In column 3 are given the numbers to be expected in pure nitrogen, calculated by proportionality from the previous results on nitrogen alone. The next column contains the difference between columns 2 and 3 and is therefore the scattering attributed to the mercury nuclei. The last column gives the numbers to be expected from Mott's theory for heavy nuclei as given below, taking into account the first term in Z in the square brackets:

$$q = \pi n t \left(\frac{Ze^2}{m_0 c^2} \right)^2 \frac{(1 - \beta^2)}{\beta^4} \left[\cot^2 \frac{\theta_1}{2} - \cot^2 \frac{\theta_2}{2} - 2\beta^2 \log_e \frac{\sin \frac{\theta_2}{2}}{\sin \frac{\theta_1}{2}} \right. \\ \left. + \frac{2\pi\beta Z}{137} \left(\sin \frac{\theta_1}{2} + \operatorname{cosec} \frac{\theta_1}{2} - \sin \frac{\theta_2}{2} - \operatorname{cosec} \frac{\theta_2}{2} \right) \right. \\ \left. + \text{extra terms in } \left(\frac{Z}{137} \right)^2, \left(\frac{Z}{137} \right)^3, \text{ etc.} \right]$$

where q = fraction scattered between angular limits θ_1 and θ_2 , Z , t and n = atomic number, number of atoms per c.c. and thickness of scatterer respectively, e and m_0 = charge and rest mass of the electron, and $\beta = v/c$ = ratio of the velocity of the β -particles to that of light. Appropriate integration has, of course, been made over the number-velocity curve of the β -ray spectrum of Ra E between $\beta = 0.86$ and 0.95 .

TABLE II. INELASTIC COLLISIONS

θ°	E_1	E_2	$(E_1 - E_2)$	$(E_1 - E_2)/E_1$ %
20	1080	973	107	10
21	738	500	238	32.5
21	973	500	473	48
24	813	500	313	39
26	600	266	334	56
32	1080	266	814	76
49	652	180	472	73
82	866	266	600	69
110	809	309	500	56

In Table II are recorded all the cases observed in which inelastic collision occurred; the lowest energy loss which could be detected with certainty was estimated at about 10 %. The angle of deflexion is denoted by θ ; E_1 and E_2 are the energies in kV before and after the collision respectively. The change

in energy was determined from the increase in curvature of the track after collision. In addition to the nine instances given in Table II, four further cases were observed in which nearly the whole of the energy of the incident β -particle disappeared. The track after the collision was short and heavily ionized and the velocity of the β -particle was estimated from its remaining range. The results are collected in Table III. The third case, corresponding to $\theta = 91^\circ$, is shown in fig. 3, Plate 2.

TABLE III. LARGE ENERGY LOSSES

θ°	E_1	E_2	$(E_1 - E_2)$	$(E_1 - E_2)/E_1$ %
$\sim 0^\circ$	1170	15	1155	99
45	575	15	560	97
91	808	90	708	89
119	1045	120	925	89

As indicated in a previous publication (Champion and Barber 1937), we have also observed what appear to be the production of low energy pairs of positrons and electrons. Examples are shown in figs. 1 and 2, Plate 2. The fast β -particle track ends suddenly, and at this point arise two low-energy electron tracks curving in opposite directions. No collisions of similar appearance were found in sixty times as much data using nitrogen alone. In Table IV, column 1 contains the initial energy of the incident β -particle in kV, while columns 2 and 3 contain the energies of the positron and electron respectively. In columns 4 and 5 are given the angles made by positron and electron with the initial direction of the incident β -particle, while column 6 contains the total angle between positron and electron.

TABLE IV. POSITRON-ELECTRON PAIRS

E_β	E_+	E_-	ϕ_+°	θ_-°	$(\phi + \theta)^\circ$
1050	150	13	48	77	125
1105	13	8	37	74	111
1105	45	4	—	—	~ 145

4. DISCUSSION

We shall first consider the inelastic collisions and the cases of pair production.

(a) *Inelastic collisions*

Referring to Table II it will be seen that nine examples were found in which a β -particle lost more than 10 % of its initial energy. From the theory of Bethe and Heitler (1934), the number of cases of sudden energy loss of

this magnitude to be expected in the present experiment was about two. The frequency with which sudden energy losses occur may therefore be considerably greater than that predicted by theory. If we include also the four very large energy losses of Table III, the experimental results show over six times the number predicted by theory. This result is in approximate agreement with the value given by Klarmann and Bothe (1936) for the inelastic scattering of β -particles of energies ranging from 0.5 to 2.5 mV by krypton and xenon nuclei. However, the values given in Table II alone show a fairly steady increase of energy loss with angle of deflexion; the inclusion of the first two cases in Table III would not be in agreement with this relation. While there may be no connexion between magnitude of energy loss and angle of deflexion, it was thought best to present the results in this form. It must be remembered that Bethe and Heitler's calculations refer to losses of energy by radiation processes, whereas the data given in Table III may refer to some other process such as the expulsion of a neutrino, or it may be connected with the repulsive field postulated later.

(b) Pair production

Referring to Table IV, which contains the data on the positron-electron pairs, in the first place it will be observed that the incident β -particle always has a kinetic energy just greater than 1 mV. This is consistent with the fact that the creation of a pair of electrons would require an available energy equal to the sum of the rest masses of the two electrons. It will be further noticed that energy is approximately conserved and that the kinetic energy of the positron is always greater than that of the electron, which is to be expected if their kinetic energies are influenced by the positive charge of the mercury nucleus. The total length of track of the incident β -particles with energy greater than 1 mV was about 90 m.; taking account of the partial pressure of the mercury dimethyl vapour, this gives an effective cross-section for pair production of about 3×10^{-22} cm.². Statistical fluctuations would clearly prevent much reliance being placed upon this figure, but it may be remarked that it is in approximate agreement with the cross-section reported by Skobelzyn and Stepanowa (1935) for the production of positrons when β -particles of energies 1–3 mV impinge on solid lead. The non-occurrence of the phenomenon in the previous experiments with pure nitrogen indicates that the probability of pair production increases fairly rapidly with the atomic number.

The absence of a second negative electron track at the point of production of the pair is of considerable interest. Its energy, after escape from the nucleus, may have been less than about 2000 V, rendering the track of the

particle very difficult to detect; but the possibility of absorption by the nucleus must not be left out of account. Unfortunately the mercury isotopes 198, 199, 200, 201, 202 and 204 are all present in comparable proportions, so predictions concerning the mass of the nucleus concerned are very uncertain. Absorption of the electron would lead to an isotope of gold. Very little appears to be known about the stable isotopes of this element, only one isotope, 197, being definitely established. However, this isotope is known to give rise to two β -active isotopes 196 and 198, under neutron bombardment and it was thought worth while to attempt to detect β -activity when mercury is bombarded with fast β -particles.

A hollow cylinder of mercury, about 3 mm. thick, contained between two concentric cylinders of thin cellophane was placed round a radon source of over 100 m.c. strength and was irradiated by the emitted β - and γ -radiation for about a day. The mercury cylinder was then placed round a tube counter of thin aluminium as in the experiments of Widdowson and Champion (1938). No activity was detected. While, of course, the most obvious conclusion is that the β -particle was not absorbed, much more elaborate experiments would be required to decide this point.*

(c) *Elastic collisions*

The most interesting result of these measurements is the great discrepancy between theory and experiment shown by the elastic scattering. Referring to Table I it is at once clear that the number of β -particles scattered between angles of 20 and 180° by mercury nuclei is only of the order of one-seventh the number predicted by Mott's theory. We are indebted to Professor Chadwick for drawing our attention to the possibility that the mercury dimethyl might have been partially removed from the chamber after a few expansions; this would, of course account directly for the reduction in the observed scattering. As such removal would change the

* Other possibilities might be (a) the isotope was not radioactive, or (b) the half-life was very long, or (c) the emitted particles were of extremely low energy, or finally (d) the half-life lay between 0.1 and 10 sec. The counter-experiments were such that a period of less than 10 sec. would have been difficult to detect, but if the period was less than 0.1 sec. the β -activity should show itself in the expansion chamber photographs as an additional track at the low-energy branch. No such track has been observed. Since Au^{198} is known to be β -active, it seems unlikely that Hg^{198} undergoes nuclear transmutation in the pair production process. However, as this isotope is present to the extent of only 10 %, its exclusion does not assist appreciably in deciding whether or not transmutation occurs at all. If stable gold is, in fact, formed, its presence would only be detected by a gradual accumulation of gold in the bombarded mercury. Experiments in the discharge tube using electrons with energies just greater than 1 mV might decide this point.

expansion ratio sufficiently to produce marked differences in the appearance of the tracks it seemed improbable that this had occurred. However, a further 250 photographs were taken with mercury dimethyl actually on the floor of the expansion chamber and exerting a pressure of not less than 4 cm. of mercury during the whole of the experiment. The results gave thirteen deflexions greater than 20° in a total track length of about 40 m. This is to be compared with our previous results of 152 deflexions in 350 m. The relative numbers are observed to be of the same order.*

These experimental results are in approximate agreement with those of Klarmann and Bothe (1936) who found less than one-fifth of the number predicted by theory for β -particles of energies ranging from 0.5 to 2.5 mV scattered by krypton and xenon nuclei at angles greater than 40° . It should perhaps be mentioned that these deviations from theory are in the reverse direction to those found experimentally by Stepanowa (1937) using nitrogen alone. Stepanowa found that the number scattered by nitrogen nuclei was about 1.5 times the theoretical for β -particles of energies 0.2–1.1 mV and about 30 times that to be expected, in the region 1.5–3 mV. It would not appear possible to explain these diverse results at present, particularly since the scattering of β -particles of energies 0.4–1.1 mV by nitrogen nuclei alone was found (Champion 1936) to be in approximate agreement with theory. Similar agreement with theory was found for the scattering of β -particles of mean energy 0.3 mV between angles of 24° and 40° by Al, Cu, Ag and Au in the early experiments of Chadwick and Mercier (1925). These results were extended to the angular range of 95° – 172° by Neher (1931) using electrons of energy 0.15 mV, when the scattering by Al was found to differ from the theoretical by only about 30 %. The *multiple* scattering of cosmic-ray electrons of energies about 100 mV would also seem to be in agreement with existing theory according to the experiments of Blackett

* It seemed to us remotely possible, however, that although the mercury dimethyl was not removed, since it was so much denser than the nitrogen present, it might take up a position at the bottom of the chamber. Fortunately, the piston-floor of the chamber was of curved rubber and its movement was not streamline. It therefore seemed probable that the motion of the piston in returning to the position preparatory to the next expansion might be sufficient to produce complete mixing. To test this we carried out some experiments with visible bromine vapour contained in glass cylinders of the same width and height as the expansion chamber. Slight movement in the lower cylinder was found sufficient to cause complete visual mixing of bromine vapour introduced initially at the bottom of the vessel. A rotating fan has now been installed inside the expansion chamber and stirs the gas before each expansion. Unfortunately we are unable to obtain a further supply of mercury dimethyl but the scattering in iodine is under investigation with the new arrangement.

and Wilson (1938). It would appear that the present position with regard to the elastic scattering of electrons by nuclei may be summarized as follows:

- (1) the scattering is normal for light elements up to about 1 mV,
- (2) the scattering is normal for heavy elements up to about 0.3 mV,
- (3) the scattering in nitrogen is much greater than that to be expected for energies in the neighbourhood of 2 mV,
- (4) the scattering by the heavy elements Kr, Xe and Hg is much less than that to be expected in the region of 1 mV,
- (5) the multiple scattering by Cu and Pb is normal for energies of the order of 100 mV (cosmic rays).

It does not seem possible to explain the phenomena presented in this paper in terms of existing theory. We suggest that the simplest interpretation of the reduced scattering is provided by the assumption of a repulsive field superposed on the ordinary nuclear Coulomb field during the scattering of fast electrons. Whatever may be the properties of such a field it would seem necessary for its effect to vary fairly rapidly either with atomic mass or atomic number since for particles of the same energy (1 mV) the scattering is found to be normal for nitrogen and anomalous for mercury. It is further clear that such a repulsive force must be confined to electron encounters since the scattering of α -particles is well known to be normal.

ACKNOWLEDGEMENTS

We wish to express our thanks to the Chemical Laboratory of the University of Cambridge for the gift of the mercury dimethyl, to the Royal Society for a grant towards part of the expansion apparatus, to the General Electric Company for suggestions concerning the illuminating lamps, to Dr Gupta for permission to use his apparatus in the counter experiments, and to Professor Ellis for valuable criticism.

We are also indebted to Mr G. F. Gainsborough for assistance in taking and examining some of the photographs.

SUMMARY

The scattering of β -particles of about 1 mV energy by mercury nuclei is examined with an expansion chamber for angles of scattering greater than 20° . The absolute number scattered is much less than would be expected from Mott's formula. There is also indication of more inelastic collisions than are predicted by the theory of Bethe and Heitler. Photographs are

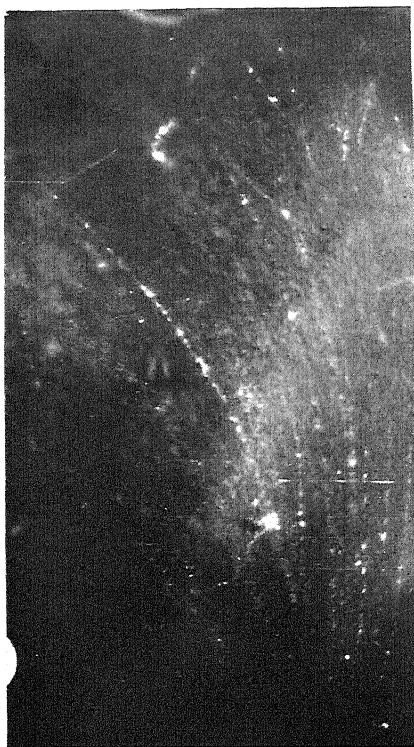


FIG. 1

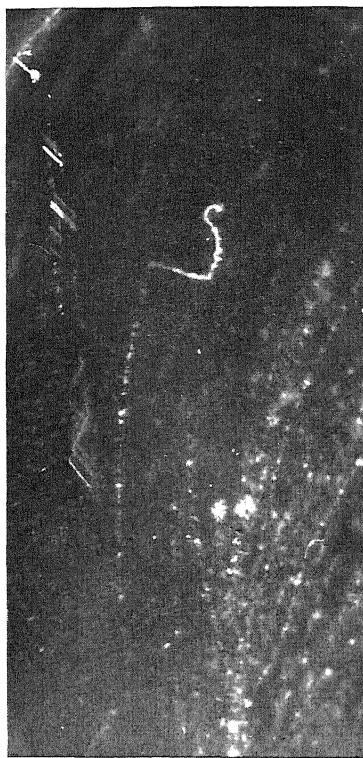
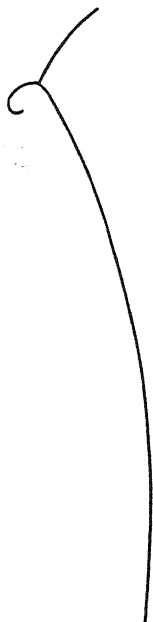


FIG. 3

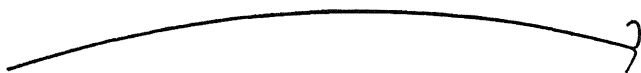


FIG. 2

reproduced which show the sudden stoppage of a fast β -particle with the simultaneous production of what appear to be low energy pairs of positrons and electrons. It is suggested that the results may be qualitatively explained on the assumption of a repulsive field between electrons and nuclei for close distances of approach.

REFERENCES

- Bethe and Heitler 1934 *Proc. Roy. Soc. A*, **146**, 83.
 Blackett and Wilson 1938 *Proc. Roy. Soc. A*, **165**, 209.
 Champion 1936 *Proc. Roy. Soc. A*, **153**, 353.
 Chadwick and Mercier 1925 *Phil. Mag.* **50**, 208.
 Champion and Barber 1937 *Nature, Lond.*, **140**, 105.
 Klarmann and Bothe 1936 *Z. Phys.* **101**, 489.
 Mott 1932 *Proc. Roy. Soc. A*, **135**, 429.
 Neher 1931 *Phys. Rev.* **38**, 1321.
 Skobelzyn and Stepanowa 1935 *J. Phys. Radium*, **6**, 1.
 Stepanowa 1937 *Phys. Z. Sowjet.* **12**, 5, 550.
 Widdowson and Champion 1938 *Proc. Phys. Soc.* **50**, 185.

DESCRIPTION OF PLATE 2 (1.5 \times natural size)

FIG. 1. The production of a low energy positron-electron pair, the data for which are given in row 1, Table IV. A diagram accompanies the photograph.

FIG. 2. Another pair, but of exceptionally low energy; the measurements are given in row 3, Table IV. Accompanied by diagram.

FIG. 3. A fast β -particle has made a highly inelastic collision and lost nearly all its energy. The measurements are given in row 3, Table III. The slow electron track is also observed to undergo a nuclear deflexion. It was assumed that there was no energy loss in this deflexion since there was no appreciable increase in ionization after the collision.

Photographic sensitivity and the reciprocity law at low temperatures

BY W. F. BERG AND K. MENDELSSOHN

(Communicated by F. A. Lindemann, F.R.S.—Received 3 June 1938)

The experiments described below were undertaken in order to examine a theory by Gurney and Mott (1938) on the formation of the latent image. The theory adopts the old conception, that the latent image consists of a speck of silver, which is produced from the silver halide by a two-stage process. The first stage consists in the lifting of an electron from a bromine ion into the conductivity band in which the electron moves until it is trapped by a region of low potential, the "sensitivity speck". The second stage consists in the movement of the positive silver ions up to the charged sensitivity speck. The ionic movement should depend on temperature as $e^{-E/kT}$ and one should therefore expect the sensitivity of photographic materials almost to disappear at sufficiently low temperatures. Experiments at liquid air temperature have been made by several investigators (see, for instance, Sheppard, Wightman, and Quirk 1934). The sensitivity found is of the order of 5–10 % of room temperature sensitivity and appears surprisingly high in view of the above considerations. It seemed desirable to lower the temperature still further by using liquid hydrogen and to test in particular for the effect of change in intensity and time, while keeping the exposure, i.e. the product of intensity with time, constant.

A great amount of work has been done on the reciprocity failure of sensitive materials, by which is meant the fact that for a constant exposure the density obtained after processing depends on the intensity level at which the exposure was given. (Webb 1933 *a, b*; Jones and Webb 1934). It appears that for every material there is an optimum rate of reception of energy. The position of the optimum depends on temperature (Webb 1935), and for this reason alone the investigation of the reciprocity failure at as low a temperature as possible seemed desirable. It was hoped that some information on the mechanism of the latent image formation could be obtained in this way.

APPARATUS

The apparatus is shown in fig. 1. The sensitive material was covered with opaque material into which ten holes had been punched; over each of these

holes was glued a piece of film exposed to a different calibrated density thus forming in effect a step-wedge.

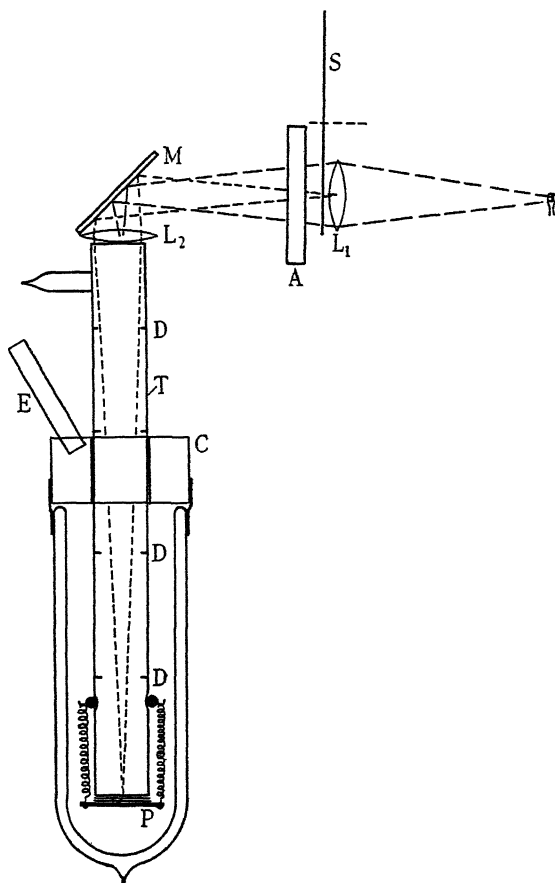


Fig. 1

The film and step-wedge were clipped to the flat bottom of a cylinder *T* made from pyrex glass, by means of a metal plate *P* attached to the cylinder by springs. The inside of the cylinder was covered with black flock paper and also contained a series of diaphragms *D* to cut out stray reflections. The top and bottom of the cylinder consisted of flat pieces of glass. The cylinder was exhausted to avoid condensation of air during the cooling with liquid hydrogen.

The cylinder was carried by a cap *C* which at the same time provided a lid for the Dewar flask containing the cooling medium. In order to make a light-tight and air-tight joint a rubber sleeve formed the connexion between the

cap and the Dewar vessel. For the liquid hydrogen experiments an exhaust pipe E leading from the cap discharged the evaporating hydrogen outside the building.

The cylinder was mounted vertically. Above it was fixed an adjustable stainless steel mirror M . The tungsten filament lamps used for exposing were mounted on carriages which ran along an optical bench of a length of 4.5 m. The variation in intensity provided by this length was not, however, big enough to cover the range desired and two lamps of different wattage always had to be used. The lamps were matched visually for colour temperature and the voltages giving this colour temperature were kept constant to within 0.1 V by manual control of series resistances. The relative intensities of the lamps were determined photographically. The exposures were at first made by means of a calibrated Compur shutter which was mounted on top of the glass tube, no lenses being used in the optical system. After a time, however, the results could not be reproduced with sufficient accuracy, owing apparently to the development of defects in the shutter. It was therefore decided to use a sector shutter S driven by a phonic wheel. The phonic wheel (supplied by Tinsley) ran at 5 rev./sec. and a geared down shaft at 1 rev. in 6 sec. The slow shaft could be used to give exposures from 5 to 1/10 sec. and the fast shaft from 1/20 to 1/2000 sec. With the fast sector it was impossible to obtain single flashes by manual control of an additional shutter A . A mechanical trip in conjunction with a Compur shutter set to 1/5 sec. served to isolate a single flash. The trip mechanism was set so that the shutter stayed open with a sufficient margin of safety for the widest sector.

For the exposure time to be defined accurately with a comparatively slow-moving sector such as this, the shutter has to move very near to the material or to a point source lamp. Since both alternatives were impracticable, a projection system of lenses was used. One lens L_1 is very close to the sector and throws an image of the source of light into the second lens L_2 . This second lens is big enough to receive the whole of the image of the source of light and in turn throws an image of the first lens and of the sector shutter on to the sensitive material to be exposed. The sector therefore acts as if situated in the plane of exposure. In such an arrangement the illumination on the film is uniform and varies as the inverse square of the distance between the lamp and the first lens. The only condition which has to be fulfilled is, that the image of the source of light thrown by the first lens must not be bigger than the second lens. If this condition is fulfilled all the light going through the first lens also goes through the second lens and through the plane of exposure.

Development was usually for 4 min. in a normal metol-hydroquinone

developer at 20° C. The films were stuck to a piece of glass and brush developed. The exposures belonging to one set of experiments were always developed together.

EXPERIMENTS

A preliminary set of experiments was done to determine the relative sensitivity of a photographic material at room temperature, liquid air and liquid hydrogen temperatures (Berg and Mendelssohn 1937). The material tested was Eastman Process film. The exposures were done through a step-wedge of a density ranging from 0 to 4, and exposure times were varied so that the steps of the wedge were fully exposed. In the curves in fig. 2 the exposure It was plotted disregarding any possible deviation from reciprocity law since the experiments were of a qualitative nature. It will be shown later that no error was introduced by this procedure.

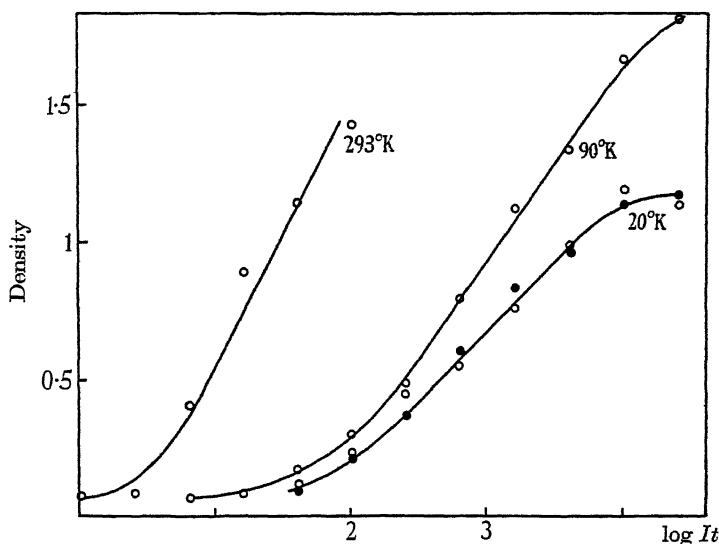


FIG. 2. Characteristic curves of process film exposed at 293, 90 and 20° K. The dots refer to an experiment in which the film has been left at low temperature for several hours after exposure.

It will be seen that for the blue sensitive Process film, the sensitivity does not drop appreciably between liquid air and liquid hydrogen temperatures. The percentage change in absolute temperature in going from liquid air (90° K) to liquid hydrogen (20° K) is about as big as that going from room temperature (300° K) to liquid air. Defining the sensitivity as the inverse

of the exposure necessary to obtain a density of 0.1 above fog we find that the sensitivity at liquid air temperature is about 7 %, at liquid hydrogen \pm % of that at room temperature for process film.

It has been suggested to us from various sides (e.g. Pohl 1937) that it is not necessary to assume that the complete process takes place at low temperatures. Possibly an electronic process only takes place at low temperature, the electrons freezing in and the ionic process then follows on warming up. Below we describe experiments which, we believe, strongly suggest this explanation to be correct.

If the explanation is correct, we should not, as at normal temperatures, find any effect of the intensity level on the developed density exposed at low temperature. The so-called reciprocity failure is explained by Gurney and Mott (1938) by recombination of electrons with halogen atoms. If at low temperatures only electronic processes take place and if the electrons freeze in at once, it seems safe to assume that the results of these processes do not depend on the rate of creation of the electrons. When the exposure is completed we are left with a certain number of negative charges, which produce the same ionic movements since the warming up takes place at the same rate for all exposures. Experiments were done to test this.

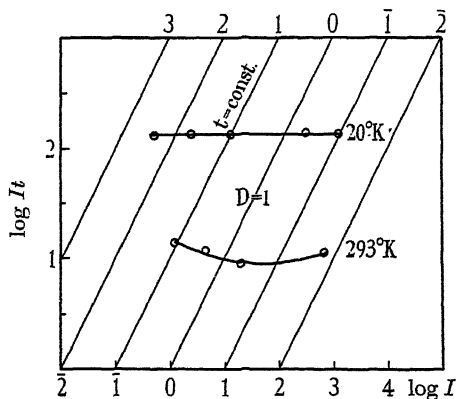


FIG. 3. Reciprocity failure diagram for process film at 20 and 293° K.

A first set of experiments was done at liquid hydrogen temperature, using a Compur shutter for timing the exposure. As a comparison, the experiments were also repeated at room temperature. The room temperature exposures show the usual bend in the reciprocity failure diagram, in which is plotted the log exposure necessary to obtain a density of 1.0 after processing, against the log intensity. At low temperature there is no reciprocity failure (fig. 3).

The procedure of arriving at the reciprocity diagrams was as follows. The

film was given an exposure at an intensity designed to expose fully through the step-wedge at the exposure time chosen. The densities of the processed film were measured and plotted against the log intensity as determined by the step-wedge and the illumination on the step-wedge. The intensity necessary to produce a density of 1.0 above fog was then found. The log exposure It necessary to produce a density of 1.0 could then be plotted against the log intensity at which the exposures were given.

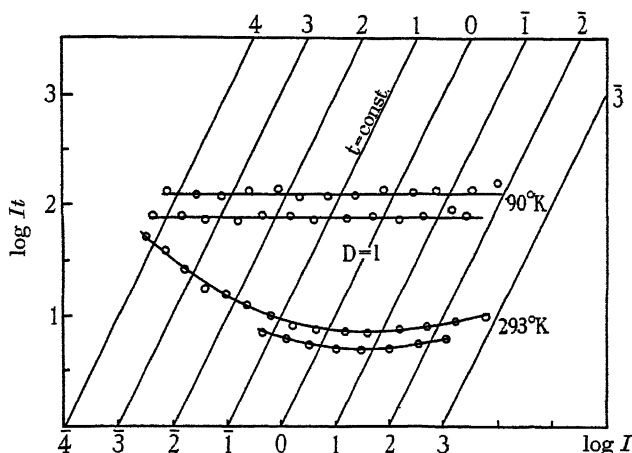


FIG. 4. Reciprocity failure diagram for process film at 90 and 293° K. The two curves at each temperature relate to different experiments.

These experiments were repeated several times with liquid oxygen using the more accurate sector shutter and the optical system described above. Here every exposure was repeated twice, in order to eliminate coating thickness variations and other unsystematic errors. The range of exposure times was also extended. The curves obtained are shown in fig. 4 and leave no doubt that within the accuracy of measurement there is no reciprocity failure at low temperature. It was not thought necessary to repeat the experiments at liquid hydrogen temperatures with higher accuracy, since the curves in fig. 2 proved that there was no appreciable difference between exposures at 90 and 20° K. Therefore no corrections need be made on the curves in fig. 2, where the exposure times had been varied and no allowances had been made for any possible reciprocity failure.

Some check experiments have been made to answer certain possible objections. The room temperature sensitivity of the materials used was found not to alter after immersion into liquid oxygen and warming up. The result of a low temperature exposure was found to be independent of the rate of

warming up within wide margins. Accidental differences in the rate of warming up will therefore not influence the results. It was also found that it made no difference whether the material was kept at low temperature for a long time after exposure or whether it was warmed up immediately. One of the curves for 20°K in fig. 2 was obtained after keeping at low temperatures for several hours, the other by warming up immediately. There is no difference between the two curves within the errors of measurement.

In the model of the formation of the latent image given by Gurney and Mott (1938) an electron of a bromine ion is first raised into an excited state. Unless the electron moves away from the positive charge represented by the bromine atom it will fall back to the normal state and be lost for the formation of the latent image. The electron must get away from the place where it is formed while at low temperature. Its energy is then essentially zero point energy and big enough to overcome the potential barrier localizing it on the silver.

The mechanism suggested above appears to provide a satisfactory explanation for the comparatively high sensitivity at low temperature. According to this mechanism an electronic process only takes place at low temperature, the electrons "freezing in" somewhere in the lattice, while the ionic movements of the silver take place on warming up. The trouble is that this picture explains too much, for if it is correct there appears to be no reason why there should be a drop in sensitivity at all. We hope to investigate this question in the near future.

The fact that photographic materials show no reciprocity failure at low temperature is, we believe, the first experimental indication that two separate processes take place during the formation of the latent image. The application of low temperature to experimental investigations thus offers some hope of further separating and studying the electronic and the ionic processes, which at normal temperatures always occur simultaneously and thus complicate matters enormously. A detailed investigation on reciprocity failure at various temperatures has been made by Webb (1935). It is apparent from his diagrams that at his lowest temperature (200°K) the ionic processes still occur simultaneously with the electronic movements, since the sensitivity still varies appreciably with the intensity level. It would seem very interesting to find out how low the temperature has to be for reciprocity failure to disappear. At 200°K the shape of the reciprocity curve seems to be essentially identical with those at room temperature, only the minimum is displaced.

It should be mentioned that under certain circumstances the fact that no reciprocity failure occurs at liquid air temperature may have valuable

applications. For Raman effect spectrograms exposure times of weeks and months are necessary and photographic materials may then be more sensitive at low temperature. The freedom from reciprocity failure may at the same time be useful for photographic-photometric work.*

Our thanks are due to Professor F. A. Lindemann, F.R.S., for his interest and his permission to have part of the apparatus constructed in his laboratory, which also supplied the liquid hydrogen, and where the preliminary experiments were done. One of us (W. F. B.) is indebted to Kodak Limited for permission to publish the work and to Mr E. R. Davies and Mr E. W. H. Selwyn of the Kodak Research Laboratory for their constant interest and advice. We also wish to thank Mr L. J. Sumner for his help with the experiments.

SUMMARY

1. Photographic materials show an appreciable sensitivity down to 20°K . Little change occurs when the temperature is dropped from 90°K to 20°K . The main difference is a drop in contrast and in maximum density.

2. No reciprocity failure occurs at temperatures of 90°K and 20°K .

3. It is suggested that at these temperatures the latent image is formed in two distinct stages: (a) an electronic process, in which the electrons are separated from the bromine ions absorbing the light, and freeze into the lattice, and (b) an ionic process, taking place on warming up, consisting of a movement of the silver and probably also the bromine atoms.

REFERENCES

- Berg, W. F. and Mendelssohn, K. 1937 *Proc. Phys. Soc.* **49**, 38.
- Gurney, R. W. and Mott, N. F. 1938 *Proc. Roy. Soc. A*, **164**, 151-67.
- Jones, L. A. and Webb, J. H. 1934 *J. Soc. Mot. Pict. Engrs*, **24**, 142-58.
- Pohl, R. 1937 *Proc. Phys. Soc.* **49**, 39.
- Sheppard, S. E., Wightman, E. P. and Quirk, R. F. 1934 *J. Phys. Chem.* **38**, 817-31.
- Webb, J. H. 1933*a* *J. Opt. Soc. Amer.* **23**, 157-69.
- 1933*b* *J. Opt. Soc. Amer.* **23**, 316-23.
- 1935 *J. Opt. Soc. Amer.* **25**, 4-23.

* The conclusions reached in this paper are fully confirmed by a paper shortly to be published by J. H. Webb and C. H. Evans. Amongst other confirmatory evidence these authors have also done experiments on reciprocity failure at 87°K and consider these to be in agreement with the experiments in this paper. We wish to thank Dr Webb for letting us see the manuscript of the paper before publication.

Experiments on the transmutation of fluorine by protons and deuterons

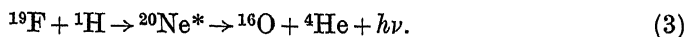
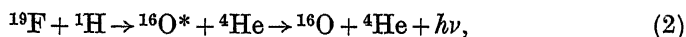
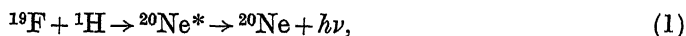
BY W. E. BURCHAM, PH.D., *Stokes Student, Pembroke College,*
AND C. L. SMITH, B.A., *Sidney Sussex College, Cambridge*

(Communicated by E. V. Appleton, F.R.S.—Received 8 June 1938)

1. INTRODUCTION

The transmutations of fluorine under bombardment with high speed particles have been studied in considerably less detail than have those of elements of lower atomic number. Of such investigations as have been made, the excitation of γ -rays by the bombardment of fluorine with protons has attracted most attention. Experiments by Hafstad, Heydenberg and Tuve (1936) and by Herb, Kerst and McKibben (1937) have, for example, established that this γ -ray excitation is a resonance phenomenon, while Gaerttner and Crane (1937) have determined the quantum energies of the radiation with some precision.

The mechanism of this γ -ray excitation has been the subject of much discussion. Gaerttner and Crane (1937) and Kalckar, Oppenheimer and Serber (1937) have in this connexion discussed the three following processes:



The known values of the γ -ray energies, however, do not satisfy any of the above equations. In the experiments to be described here a careful search has been made for groups of α -particles which, according to the processes (2) and (3), might be associated with the γ -ray emission. The emission of a homogeneous group of α -particles from fluorine under bombardment with protons of 1.69×10^6 eV energy has been reported by Henderson, Livingston and Lawrence (1934). The range of this group was 6.95 cm. and corresponded to the total energy release in (3) assuming the formation of ^{16}O in the ground state. No other groups of α -particles have been reported. The ranges of the groups of α -particles which would be formed according to processes (2) and (3) and the known values of the γ -ray energies are 0.9 and 1.8 cm. We have been unable to find any evidence for such groups, and it appears

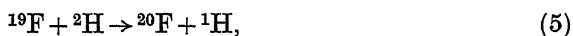
that no emission of α -particles of appreciable energy can be associated with the γ -ray emission. This result is in accordance with the resonance character of the γ -ray excitation and the apparent homogeneity of the γ -rays, and it appears to us that the γ -ray emission must occur as a result of successive radiative transitions of excited ^{20}Ne nuclei according to equation (1). Further support for this view has been obtained by a careful measurement of the energy of the group of α -particles discovered by Henderson, Livingston and Lawrence.

The present research has also included the investigation of the transmutations of fluorine under deuteron bombardment, a subject which has not previously been studied in detail. Lewis, Livingston and Lawrence (1933) reported the emission of a group of α -particles of range 3.8 cm. at 1.3×10^6 eV bombarding energy which they attributed to the process



The range to be expected according to the mass values of the particles concerned is 9.1 cm. but no such group was observed by these workers. The experiments described in this paper establish the emission of such a group together with several other groups of α -particles, the energies of which permit the construction of a nuclear level system for ^{17}O .

The radioactivity induced in fluorine under deuteron bombardment, which is supposed to be due to the formation of ^{20}F in accordance with the process



has been studied by Crane, Delsasso, Fowler and Lauritsen (1935). These workers determined the period of the radioactivity (12 sec.) and the upper limit of the β -ray spectrum (5.2×10^6 eV). Fowler, Delsasso and Lauritsen (1936) reported the observation of a group of protons of range 10 cm. which they associated with the assumed process. The energy balance in the reaction is not satisfied by these values, about 4×10^6 eV being lacking in the total energy release. We have made a careful search for groups of protons of higher energy and although one such group was found the energy balance remains unsatisfied. No protons of 10 cm. range could be detected other than a group due to contamination of the fluoride targets with oxygen. This group of protons was produced according to the process



which was studied by Cockcroft and Lewis (1936). A search was also made for γ -radiation accompanying the β -particle emission from ^{20}F , and this

was detected with Geiger counters and the identity of the decay periods of the β -particle activity and γ -ray activity established. The determination of the energy spectrum of these γ -rays will form the subject of a future paper.

2. EXPERIMENTAL METHOD

The experiments were carried out in the new high-voltage building of the Cavendish Laboratory. Protons and deuterons issuing from a discharge tube of the type described by Oliphant and Rutherford (1933) were accelerated in a three-stage vacuum tube connected to a Philips 1.2×10^6 eV D.C. generator. With this apparatus targets of about 1 sq. cm. in area could be bombarded with a magnetically resolved ion beam of about $50 \mu\text{A}$ and 1.1×10^6 eV energy, but in most of the present experiments an ion beam of $30 \mu\text{A}$ and 0.85×10^6 eV energy provided a yield of disintegration products as great as could conveniently be counted. The energy of the bombarding ions could be measured with an error of less than 1 %. A detailed description of the accelerating tube and its auxiliary apparatus will be published later.

At the higher bombarding voltages used in these experiments the disintegration of the usual surface contaminants (carbon, oxygen and deuterium) of the targets tends to obscure the presence of disintegration products of heavier elements, while the increased liberation of heat due to the impact of the bombarding ions may cause evaporation of the targets unless they are efficiently cooled. The difficulty due to oxide contamination of the target was partially removed by gold-plating all surfaces of the target holder which were exposed to the ion beam. Metallic fluoride deposits (KF , NaF , AgF , BaF_2) were prepared upon gold-plated copper discs either by evaporation of aqueous solutions of the salts or by evaporation of the fused salts and condensation of the vapours upon the disks. Three such target disks were rigidly fastened to a water-cooled copper block mounted upon a ground joint, rotation of which enabled any one of the disks to be brought into the path of the ion beam. Evaporation of the targets was greatly reduced by providing good thermal contact between the disks and the copper block, but fortuitous variations of the target efficiency could not entirely be prevented. The particles emitted from the bombarded target emerged from the vacuum through a mica window mounted upon a collimating tube, the axis of which was at right angles to the path of the incident ions. The angular spread of the emergent particles was about 10° . These particles were detected with a differential ionization chamber and were counted by means of a linear amplifier and "scale-of-two" valve counter

of the type described by Lewis (1937). The use of a cathode-ray oscillograph for visual observation of the impulses arriving at the counter has been of great convenience throughout the experiments. The range of the particles recorded by the counter was usually varied by inserting mica screens between the mica window and the ionization chamber. Where, however, the presence of two or more groups of particles of about the same range was suspected it was found to be more convenient to provide a succession of small changes of absorption by the movement of the ionization chamber. In such movements the distance between the ionization chamber and the mica window was always less than the value at which the particles might have failed to enter the chamber on account of the divergence of the beam. No correction was needed, therefore, for the variation of the solid angle subtended by the chamber at the target.

The numbers of disintegration particles (per min. per μA of target current) which were recorded by the counter for different amounts of the absorption in the path of the particles provided the data from which the distribution in energy of the products of the transmutation could be obtained. Using this method, however, fortuitous variations of the target efficiency which occurred during bombardment often gave rise to a lack of reproducibility of the results which was very troublesome when the distribution in range of groups of particles of about the same energy was being studied. To overcome this difficulty a separate ionization chamber which scanned the target through a second mica window on the target tube was used to determine the number of disintegration particles having a fixed value of the range. The absorber in front of this chamber was kept fixed during an experiment at a value such that the number of particles which were recorded was almost entirely due to a known group of particles arising from the transmutation of fluorine. The ratio of the counts of the two chambers was independent of variations of the target efficiency, and the distribution in range of the products of the transmutation of fluorine was obtained from the values of this ratio when the amount of absorber in front of the first chamber was varied over the region under investigation. Precision measurements of the range of any group of particles was usually made by direct comparison with the range of the 8 cm. group of α -particles emitted from lithium under proton bombardment using similar geometrical conditions of counting.† The difference in range of the two groups was therefore equal to the difference of absorption in front of the chamber in the

† The extrapolated range of this group of particles was taken as 8.40 cm. for an energy of the bombarding protons of 0.19×10^6 eV (Oliphant, Kempton and Rutherford 1935).

two cases, a small correction being made for the variation of the straggling with range. The nature of the disintegration particles was decided by visual observation of the size of the impulses on the oscillograph screen and by measurement of the "cut-off bias" required to prevent the impulses from being recorded by the valve counter. In this way it was easy to distinguish α -particles from protons providing that the total number of impulses per minute was not so large as to result in appreciable superposition. This latter condition was always fulfilled in an experiment by using a diaphragm of the necessary size in the path of the disintegration particles.

3. EXPERIMENTAL RESULTS

(a) Proton bombardment

The distribution in range of the α -particles which are emitted from fluorine under bombardment with protons of 0.85×10^6 eV energy is shown in fig. 1. One group of α -particles only was detected, from different fluoride targets, of mean range 5.90 cm. A search was made for α -particles of shorter range without success. With a bombarding energy of 0.85×10^6 eV the detection of range less than 1.8 cm. was impossible on account of protons

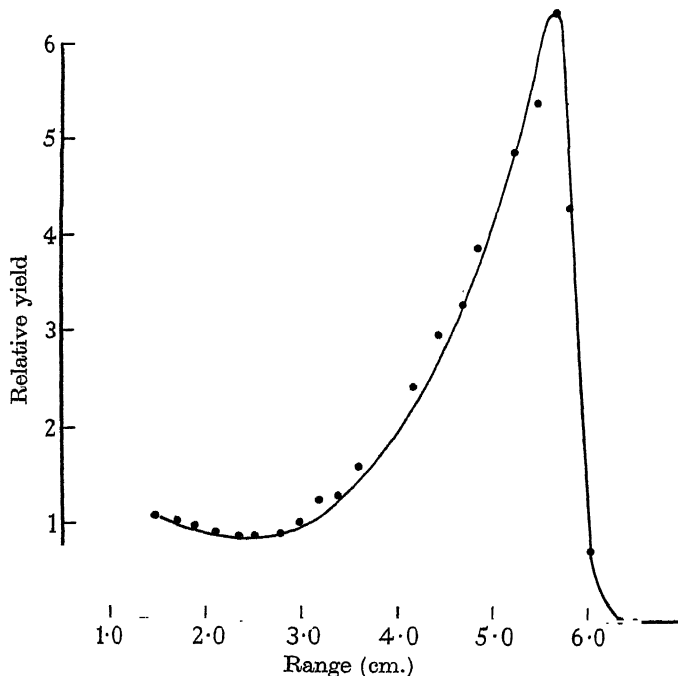


FIG. 1

scattered from the beam. This lower limit was reduced to less than 1 cm. by working at lower voltages, but no evidence of any other group of α -particles, definitely attributable to fluorine, could be obtained. The group of α -particles at 5.90 cm. showed no fine structure.

(b) Deuteron bombardment. Emission of α -particles

The results obtained by the bombardment of fluoride targets with deuterons of 0.85×10^6 eV energy were much more complex and are represented in fig. 2.

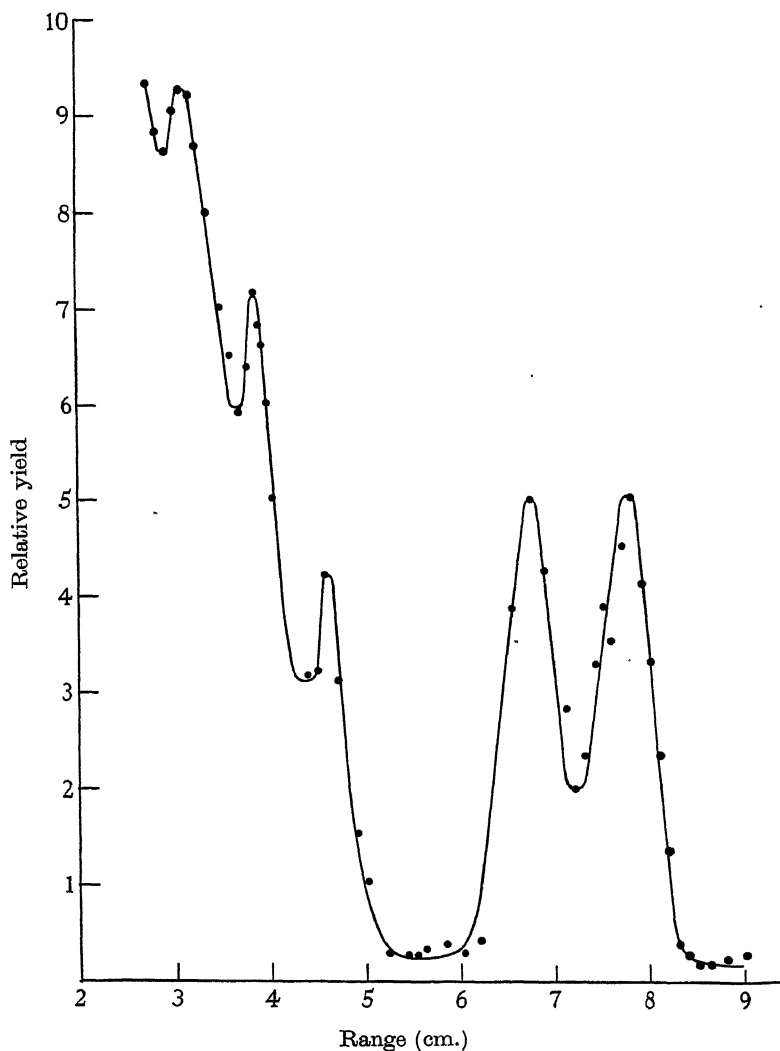


FIG. 2

Five groups of α -particles of mean ranges 8.20, 7.21, 4.84, 4.05 and 3.42 cm. have been observed, in about the same relative intensities from each of the fluoride targets used. The possibility of carbon and oxygen contamination could, however, in no case be excluded, and in order to determine whether in fact these groups were due to the transmutation of fluorine a direct comparison was made between the number-range distributions obtained from targets of KF, carbon, and KOH. The results of this comparison are shown in fig. 3.

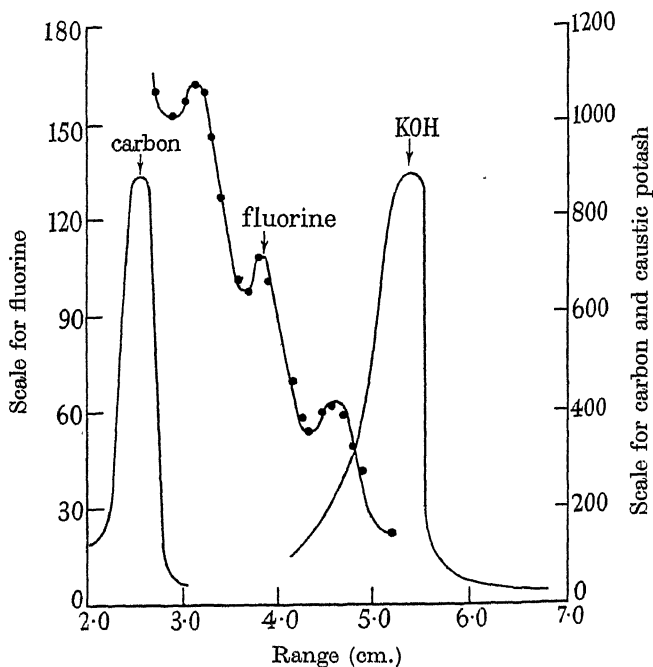


FIG. 3

The known α -particle group from carbon of mean range 2.8 cm. and the peak due to the effect of superposition of protons of 5.6 cm. range arising from the transmutation of oxygen may be seen to differ in range from the values of either of the three shortest of the five α -particle groups obtained from the fluoride targets. No particles were detected from the carbon and KOH targets in the neighbourhood of the groups of 8.20 and 7.21 cm. range, and it may therefore be definitely concluded that the five groups of α -particles obtained from fluoride targets were due to the transmutation of fluorine.

Fig. 4 shows the number-range distributions relating to the four α -particle groups of highest energy for different energies of the deuterons

incident upon thin targets. From these curves the excitation function for the groups has been determined and is shown in fig. 5. This curve shows no resonance features, and the variation of range with bombarding energy is in approximate agreement with the explanation of their origin given in § 4*b*. The possibility of the existence of other α -particle groups of even shorter range arising from the transmutation of fluorine cannot be excluded but their detection would present great difficulty owing to contamination of the targets with carbon.

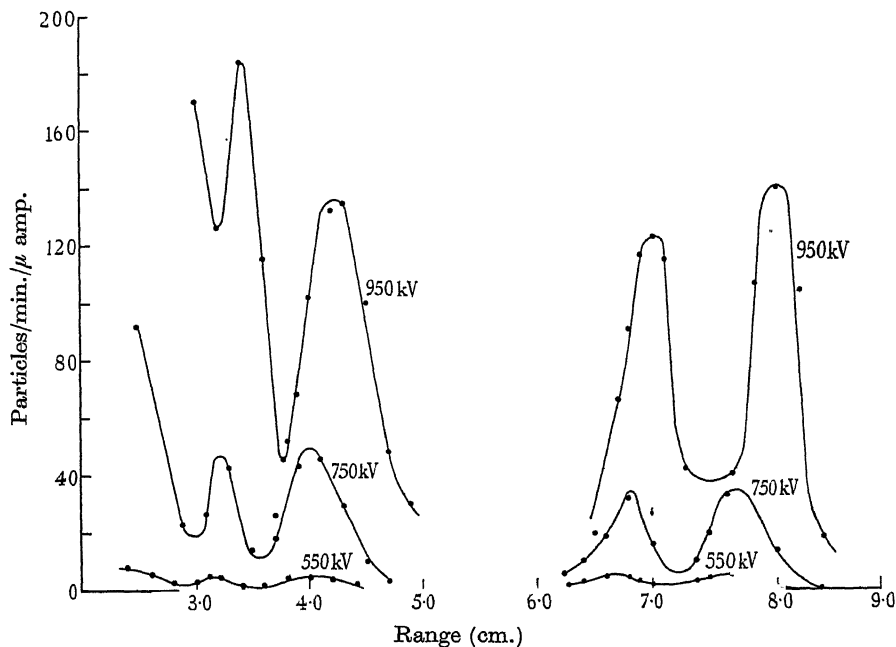


FIG. 4

(c) *Deuteron bombardment. Proton emission and production of radioactive fluorine*

A typical number-range curve for the particles emitted from fluoride targets when the bias of the counter was adjusted so that protons were recorded, is shown in fig. 6, curve *A*.

The α -particle groups of 8.20 and 7.21 cm. range are evident as is also the proton peak at 15 cm. which was proved to be due to deuterium contamination of the target. A group of protons of mean range 10 cm., as reported by Fowler, Delsasso and Lauritsen (1936), is also evident but subsequent experiments proved that both this peak and that occurring at a range of 5 cm. could also be obtained from blank targets with an intensity

which appeared to depend upon the cleanliness and particularly upon the state of oxidation of the surface. Direct comparison of the positions and relative intensities of these two groups arising from fluoride targets and similar groups obtained with oxide targets revealed a complete identity.

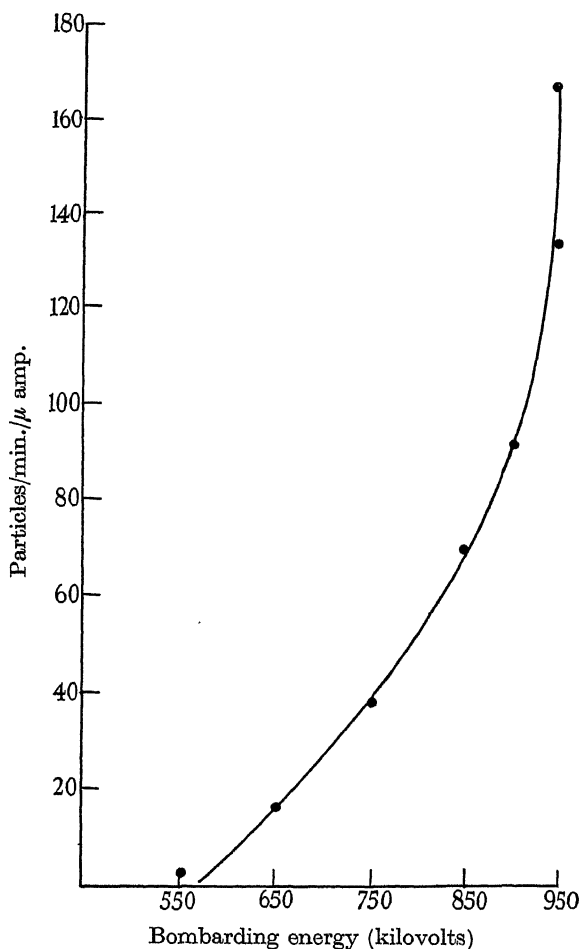


FIG. 5

It was concluded that these groups were, in every case, due to the proton emission from oxygen bombarded by deuterons (Cockcroft and Lewis 1936), and that there was no detectable intensity of protons of range about 10 cm. arising in the transmutation of fluorine. These results suggest that the group reported by Fowler, Delsasso and Lauritsen was also due to contamination of the target with oxygen.

The rather ill-defined group of protons in the region from 10.5 to 13.5 cm. appeared, however, to be due to the transmutation of fluorine. The number-range distribution obtained from the deuteron bombardment of an oxide target is also shown (curve *B*) on fig. 6. The presence of a broad group of protons attributable to the transmutation of fluorine and having a mean range of 11.6 cm. is established, but the structure of this group cannot be deduced with certainty from the results owing to the difficulty of estimating the amount of superposition due to the adjacent contamination groups.

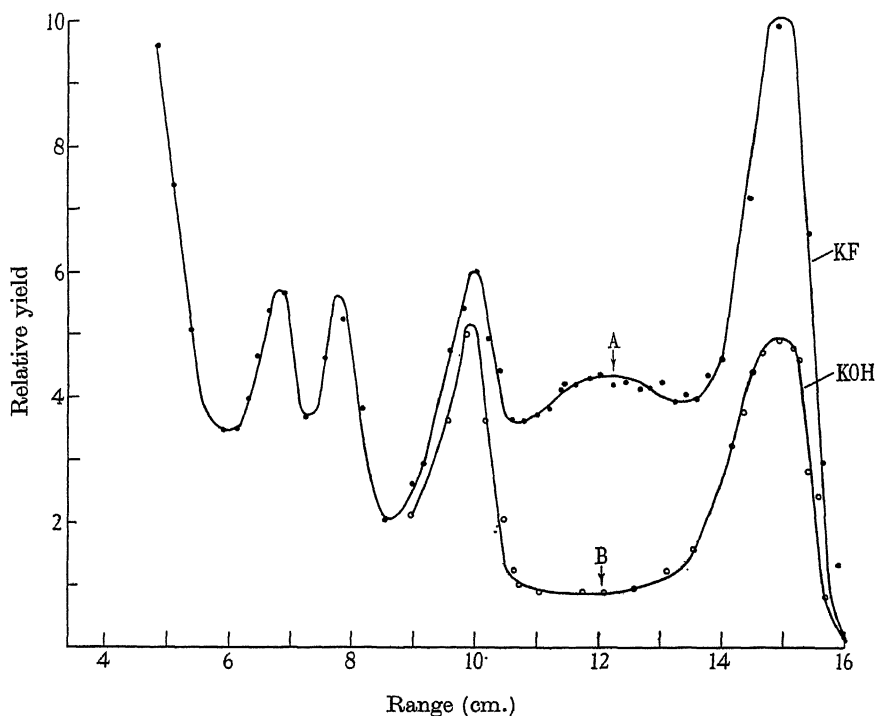


FIG. 6

The period of the β -particle radioactivity resulting from the bombardment of fluorine with deuterons has been measured and found to be 12.4 sec. in good agreement with the value (12 sec.) given by Fowler, Delsasso and Lauritsen. A direct comparison of the emission of protons of range 11.6 cm. with the radioactive emission of β -particles of 12.4 sec. period has also been attempted, and it was found that the rate of emission of the protons was about five times too small to account for the rate of formation of the radioactive product. A search was therefore made for other proton groups which might have been associated with the radioactivity, but no evidence of such

a group could be obtained in the region of from 2 to 100 cm. range. The possible existence of a group of range less than 2 cm. cannot be excluded. Approximate values of the relative intensities of the different α -particle groups are given in the following table:

Mean range of group at 850 kV	Relative intensity
8.20	1.0
7.21	1.0
4.84	1.1
4.05	1.1
3.42	2.3

4. INTERPRETATION

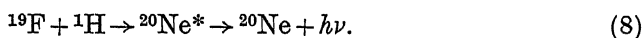
(a) *Proton bombardment*

The energy release in the process



may be calculated from the energy of the observed group of α -particles which has a mean range of 5.90 cm. at 0.85×10^6 eV bombarding energy. The value thus obtained is 7.95×10^6 eV and the mass of ^{19}F obtained by substituting the other mass values is 19.0043 in good agreement with the mass spectroscopic value 19.0045 obtained by Aston.

It appears that no α -particle group can be connected with the strong γ -radiation arising from this bombardment. The most natural explanation of this γ -ray emission, in view of the resonance features of its excitation, appears to lie in the process



Since, however, the total energy which should then appear in the γ -radiation according to this process is 13.0×10^6 eV at 400 kV bombarding energy, whereas the observed γ -radiation consists of two lines, of approximately equal intensity, having energies 5.7 and 4.0×10^6 eV, Gaerttner and Crane (1937) have suggested that the γ -emission arises in the process



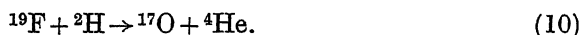
The absence of appreciable breadth in the observed lines was explained by substituting the sum of the γ -ray energies (9.7×10^6 eV) in (9) and deducing that the $(^{16}\text{O} + ^4\text{He})$ would be formed with nearly zero kinetic energy.

The value 7.95×10^6 eV which we have obtained for the energy release in the process (7) does not support their hypothesis since the *maximum* total energy which should appear in the γ -emission at a bombarding energy of 0.40×10^6 eV is 8.35×10^6 eV, differing markedly from the measured values $(5.7 + 4.0) \times 10^6$ eV of the γ -ray energies.

It seems more likely that these γ -rays arise from successive radiative transitions of ^{20}Ne and that the disagreement between the total energy release in the process (8) and the observed γ -ray energies may be reconciled by a further study of the γ -ray emission and the discovery of additional γ -rays, possibly with smaller energies than the lowest which have so far been detected.

(b) *Deuteron bombardment. Emission of α -particles*

The several groups of α -particles observed in the bombardment of fluorine with deuterons are most simply interpreted in terms of levels of excitation in the ^{17}O nucleus which is formed according to the process



The values of the energy release in this process which result from the ranges given in § 3*b* are 9.84, 9.01, 6.89, 6.07 and 5.35×10^6 eV.

Assuming that the value 9.84×10^6 eV corresponds to the formation of ^{17}O in the ground state and using our value for the mass of ^{19}F we obtain the value 17.0045 for the mass of ^{17}O , in exact agreement with the value obtained by Cockcroft and Lewis (1936) from a study of the process



The other values of the energy release in (10) must be attributed to excited states in the ^{17}O nucleus at energies 0.83, 2.95, 3.77 and 4.49×10^6 eV above the ground state. It is interesting to note that two groups of protons were observed by Cockcroft and Lewis and that the group of smaller energy may also be explained on the assumption of an excited state of ^{17}O at 0.83×10^6 eV. Evidence of the existence of excited states of ^{17}O has also been obtained from observations of the protons which are emitted in the bombardment of nitrogen with α -particles but the data is not sufficiently accurate to justify comparison with the present results. A study of the disintegration of neon by neutrons according to the process



has been made by Jaeckel who found energy losses grouped closely around

the two values 0.7 and 5.3×10^6 eV. The difference of these values may be attributed to an excited state of ^{17}O at 4.6×10^6 eV energy in good agreement with the highest excitation level deduced from our measurements.

(c) *Deuteron bombardment. Emission of protons*

It is probable that the proton group having a mean range of 11.6 cm. arises in the process leading to the formation of radioactive ^{20}F .



As stated in § 3c the number of protons in this group is insufficient to account for the amount of radioactive fluorine which is produced and it is probable that frequently the ^{20}F nucleus is formed in an excited state and that a proton of smaller energy is emitted. We may assume however that the total energy release in the process (13) may be obtained from the energy of the group of 11.6 cm. range since no protons of greater range have been observed to result from the transmutation of fluorine. The mass of ^{20}F deduced in this manner is 20.0087 . The end-point energy of the β -ray spectrum resulting from the subsequent decay of ^{20}F according to the process



is 5.2×10^6 eV whereas the value of the energy release in (14) which is obtained by the substitution of our value for the mass of ^{20}F and Aston's value for the mass of ^{20}Ne is 9.2×10^6 eV. It seems probable therefore that the ^{20}Ne nucleus resulting from the β -decay has an energy of excitation of 4.0×10^6 eV. We have obtained support for this view by some experiments which have established the radioactive emission of γ -radiation from fluoride targets after bombardment with deuterons. The half period value of decay of this γ -emission has been found to be 12.4 sec. in agreement with the half period value of decay of the β -emission. The energy of this γ -radiation is being investigated by Bower and preliminary results indicate that the radiation consists of several lines with quantum energies between 1.0 and 2.4×10^6 eV. It is of course possible that the de-excitation of the ^{20}Ne nucleus takes place in a succession of transitions. The results are not yet sufficiently precise to justify the construction of a level scheme for ^{20}Ne , but it is interesting to note that the existence of levels of excitation of ^{20}Ne in the region from 0 to 4×10^6 eV above the ground state are also required to explain the missing energy in the γ -ray emission from ^{20}Ne formed by the capture of protons by ^{19}F (see § 4a).

We welcome this opportunity to express our indebtedness to Professor M. L. Oliphant who was mainly responsible for the design and installation of the accelerating tube and other experimental equipment of the Cavendish high voltage laboratory. We regret that, owing to his appointment to the Chair of Physics at Birmingham University, he was unable to take part in the experimental researches which have been carried out with this apparatus. We wish also to thank Mr P. I. Dee for his interest and collaboration throughout the whole of this work, and to acknowledge the help which we have received from Dr W. B. Lewis in the design of improved counting methods.

SUMMARY

The transmutations of fluorine under bombardment with artificially accelerated protons and deuterons have been systematically investigated.

The energy of the group of α -particles which is emitted during bombardment of fluorine with protons has been accurately determined and leads to a value 19.0043 for the mass of ^{19}F , in good agreement with the mass spectroscopic value. From the value of the energy release in this process and from the absence of α -particle groups of smaller energy it is concluded that the γ -ray emission from fluorine under proton bombardment arises from successive transitions of excited ^{20}Ne nuclei.

Five groups of α -particles were found to be emitted during bombardment of fluorine with deuterons and the energies of these groups have been determined. From the energy values of these groups it is concluded that excitation levels at 0.83, 2.95, 3.77 and 4.99×10^6 eV above the ground state exist in ^{17}O nuclei.

The emission of a group of protons which is believed to be associated with the formation of ^{20}F has been observed. From the energy of these protons the value 20.0087 has been deduced for the mass of radioactive ^{20}F . An energy release of 9.2×10^6 eV is therefore to be expected in the radioactive decay process $^{20}\text{F} \rightarrow ^{20}\text{Ne} + e^-$ and the disagreement between this value and the observed end point energy of the β -rays (5.2×10^6 eV) has been partially reconciled by the observation that a radioactive emission of γ -radiation accompanies the β -emission and decays with the same period.

REFERENCES

- Cockcroft and Lewis 1936 *Proc. Roy. Soc. A*, **154**, 261.
 Crane, Delsasso, Fowler and Lauritsen 1935 *Phys. Rev.* **47**, 971.
 Fowler, Delsasso and Lauritsen 1936 *Phys. Rev.* **49**, 561.

- Gaerttner and Crane 1937 *Phys. Rev.* **52**, 582.
Hafstad, Heydenberg and Tuve 1936 *Phys. Rev.* **49**, 866.
Henderson, Livingston and Lawrence 1934 *Phys. Rev.* **46**, 38.
Herb, Kerst and McKibben 1937 *Phys. Rev.* **51**, 691.
Kalekar, Oppenheimer and Serber 1937 *Phys. Rev.* **52**, 279.
Lewis 1937 *Proc. Camb. Phil. Soc.* **33**, 549.
Lewis, Livingston and Lawrence 1933 *Phys. Rev.* **44**, 55.
Oliphant, Kempton and Rutherford 1935 *Proc. Roy. Soc. A*, **149**, 406.
Oliphant and Rutherford 1933 *Proc. Roy. Soc. A*, **141**, 259.
-

The thermal properties and heats of adsorption of films on vitreous silica

BY W. G. PALMER

St John's College, Cambridge

(Communicated by Sir William Pope, F.R.S.—Received 14 June 1938)

As a contribution to the few existing data of sufficient accuracy for the calculation of heats of adsorption and other thermal properties of films on adsorbents simpler than charcoal, a series of measurements has been made of the adsorption of the vapours of benzene, acetone, and methyl alcohol on vitreous silica at temperatures between 70 and 25° C. Previous experience with this adsorbent (Palmer and Clark 1935; Palmer 1937) showed that it was well suited for the purpose, particularly in respect to the very rapid attainment of equilibrium. Complete isothermals at 25° for these vapours up to saturation pressures have already been investigated and recorded (Palmer 1937). The use of vapours rather than gases not only admits of experimental simplicity, but has the further advantage that the free energy of adsorption (or adsorption potential) is immediately calculable from the relative pressure at equilibrium. It is hoped that the selection of substances is representative, in that they are respectively non-polar but polarizable, polar but non-associated, polar and associated. These molecular distinctions appear prominently in the properties of the adsorption films.

EXPERIMENTAL

The preparation of the silica adsorbent, manipulation, and methods of measurement were in principle the same as those previously described, with the following modifications. The volumetric portion of the

apparatus, including the manometer bulbs b_1 and b_2 (fig. 1; cf. Palmer 1937, fig. 1) was housed in an air thermostat T_1 regulated to 42.5° , and the horizontal adsorption tube A immersed in a water thermostat T_2 that could be set at temperatures ranging from 70 to 20° . Connecting tubes were kept at 43° by means of electrically heated wrapping wire. Since there was a maximum temperature difference of only about 20° along the tube t (of 1 mm. bore) and a pressure of vapour never less than about 1 mm., no correction for thermal pressure difference was necessary (see also below, p. 194). The powdered adsorbent weighed about 15 g. and exposed about 7 sq. m. of surface; its exact specific surface was determined by the rate of solution in hydrofluoric acid, as previously described (Palmer and Clark 1935).*

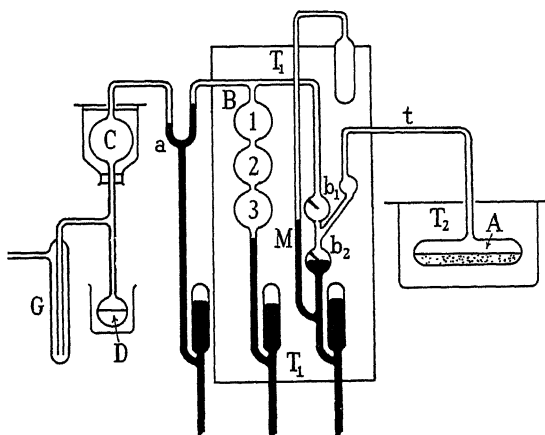


Fig. 1. Apparatus.

The vapour before expansion into the adsorption vessel A was held in the calibrated bulb system B 1, 2 and 3 (of total capacity 40 c.c.), of which one or more of the sections was brought into use depending on the final equilibrium pressure desired: the pressure was read on the manometer M with mercury set to the point b_1 . After expansion into A , the final pressure was read with mercury at b_2 , thus avoiding the disturbance of the temperature equilibrium in the connecting tube and in A , which would take place if mercury were raised again to b_1 before the equilibrium pressure was ascertained. The vessel A was then locked off by raising the mercury to b_1 , and the pressure, now that in the volumenometer, again read before the admission of a further portion of vapour.

As isothermals for different temperatures were to be used to calculate

* The author is indebted to Mrs W. G. Palmer for estimating the specific surface.

the temperature variation of the film properties and the heat of adsorption, it was clearly of crucial importance that the general surface properties, especially the area, should be invariable or that correction should be readily applicable for any changes. To ensure this two alternative methods of experimenting were adopted:

(a) *A bracketing method.* An isothermal at 70° was determined, the last and highest pressure being reached by flooding all the bulbs *B* with mercury. The temperature of the thermostat T_2 was then slowly lowered to 60° and held at that temperature for at least an hour, after which the new and slightly lower pressure was read. The bulbs were then emptied and another point of the 60° isothermal obtained from a second still lower pressure reading. The apparatus was then thoroughly evacuated and the silica heated at 100° *in vacuo* for 2 hr. to remove the last traces of adsorbate, which were collected in *G* by cooling with liquid nitrogen. The isothermal for 50° was next determined, the temperature at the end being first raised to 60° and then lowered to 40° ; two points at each of these temperatures were obtained in the above manner, with confirmation of the 50° points between. After evacuation and heating to 100° as before, the isothermal for 25° was finally determined, giving on rise to 40° at the end two points at that temperature. Sets of experiments were rejected in which the "bracketing" at 40° and 60° was not within 3 %. It was found inadvisable to attempt to secure full isothermals at more than three temperatures.

(b) *A method of continuous temperature decrease.* Isothermals for 70° and 30° were first determined in separate experiments. In a third series the silica was first brought to 70° and three points taken at higher pressures to confirm the earlier measurement. The temperature was then lowered to 30° in stages of 10° , being held for at least half an hour at each stage, during which interval two points were registered as in method (a). As the temperature fell, measured amounts of vapour were removed from the system by cooling out with liquid nitrogen, so as to maintain an approximately constant range of relative pressure throughout the series. Finally, the points at 30° confirmed those already found on the full isothermal. Such methods are possible only in adsorption systems in which equilibrium is very rapidly established after change of conditions. Fig. 2 depicts the results of method (b) applied to acetone, by showing the adsorption potential ϕ plotted against the adsorption.

Before adsorption measurements were begun in any series the silica was brought to the required temperature in an atmosphere of about 100 mm. of nitrogen, which was pumped off before admission of vapour. When heated *in vacuo* the badly conducting powder only slowly reaches uniform tem-

perature, and in the absence of precautions such as the above procedure, quite misleading adsorption data may be obtained, especially at the higher temperatures. No detectable adsorption of nitrogen takes place in the temperature range of this work. In using the methods (a) and (b) the pressure of vapour must not fall below about 30 mm. if uniform temperature is to be maintained.

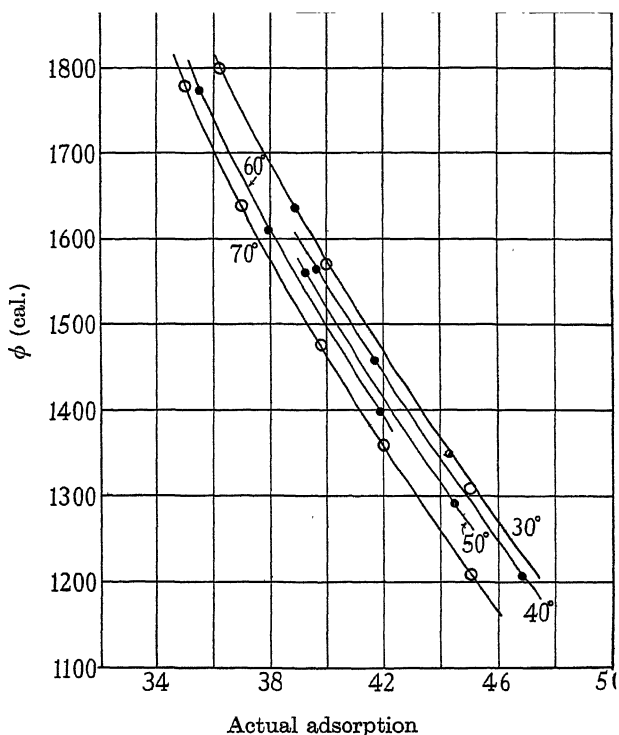


Fig. 2. Acetone.

As the apparatus did not permit the use of higher pressures at 70° than correspond to one layer, the whole work was restricted to this range of adsorption, but the nature of the results obtained leaves little doubt that they can be confidently extrapolated to larger adsorptions. A standard temperature treatment at 100–110° was adhered to in preparing and evacuating the silica; the use of higher temperatures tends to cause some drift in the surface properties. Adsorption potentials for a given adsorption are all rather higher than in the previous work wherein temperatures of about 200° were used. The effective “dead-space” was determined by blank experiments with nitrogen, the volumometer being maintained at 42°

and the adsorption tube at 70° and at 25°; no difference was detectable in the two estimates for these temperatures (which were the extremes used), thus showing that thermal effusion effects were negligible. Special experiments made to estimate the divergence of the vapour of methyl alcohol from the ideal at 42° showed that only very small corrections were involved; it had been shown previously (Palmer 1937) that even at 25° and much higher pressures the corrections for benzene and acetone were negligible.

Table I summarizes the experimental observations. In this table p is the equilibrium pressure in mm. of mercury, ϕ the adsorption potential, and a the adsorption in micromols per sq. m. $\phi = RT \log p_s/p$.

TABLE I

BENZENE. (Actual surface of silica 7.1 sq. m.)

70°: $p_s = 547.4$ mm.			50°: $p_s = 269.0$ mm.			25°: $p_s = 94.0$ mm.		
p	$\log_{10} \phi$	a	p	$\log_{10} \phi$	a	p	$\log_{10} \phi$	a
6.5	3.4835	0.29	3.1	3.4604	0.54	2.5	3.3353	1.34
15.7	3.3872	0.84	7.6	3.3630	0.96	5.6	3.2261	2.25
29.1	3.3043	1.55	15.6	3.2653	1.97	10.1	3.1243	3.18
44.7	3.2358	2.23	29.5	3.1553	2.89	16.9	3.0103	4.26
59.3	3.1837	2.66	42.6	3.0761	3.60	24.6	2.9030	5.23
68.65	3.1541	2.96	63.6	2.9698	4.65	33.6	2.7881	6.30
104.4	3.0562	3.77	94.6	2.8291	5.93	47.2	2.6140	7.99

ACETONE. (Actual surface of silica 7.1 sq. m.)

70°: $p_s = 1190$ mm.			30°: $p_s = 282.7$ mm.		
p	$\log_{10} \phi$	a	p	$\log_{10} \phi$	a
4.1	3.5900	2.05	0.2	3.6432	1.46
13.1	3.4909	2.90	1.6	3.4965	2.96
28.1	3.4094	3.60	5.5	3.3777	4.00
44.0	3.3550	4.05	12.5	3.2762	4.91
71.2	3.2865	4.57	20.75	3.1992	5.65
103.5	3.2244	5.07	27.6	3.1486	6.09
131.6	3.1796	5.65	38.6	3.0808	6.67

METHYL ALCOHOL. (Actual surface of silica 7.2 sq. m.)

70°: $p_s = 927$ mm.			50°: $p_s = 406$ mm.			25°: $p_s = 125$ mm.		
p	$\log_{10} \phi$	a	p	$\log_{10} \phi$	a	p	$\log_{10} \phi$	a
6.3	3.5352	1.21	5.1	3.4520	2.18	1.4	3.4282	2.50
18.1	3.4323	2.23	15.4	3.3256	3.51	5.1	3.2808	4.00
46.9	3.3118	3.40	31.1	3.2206	4.51	11.15	3.1592	5.18
75.2	3.2369	4.16	55.0	3.1116	5.60	20.75	3.0296	6.45
90.8	3.2028	4.52	69.5	3.0575	6.12	28.2	2.9487	7.22
138.5	3.1159	5.36	103.85	2.9453	7.23	41.5	2.8182	8.60

DISCUSSION

The following symbols will be used, in addition to those already mentioned in connexion with Table I:

- θ = fraction of surface covered.
 A = area per molecule (Angstroms).
 (A) = fictitious area per mol. for film of more than one layer.
 s = slope of $\log_{10} \phi$, a lines.
 F = lateral force (dynes cm.).
 α = thermal expansion of film $[\partial \log A / \partial T]$.
 α' = thermal expansion of bulk phase $[\partial \log V / \partial T]$.
 β = compressibility of film $[\partial \log A / \partial P]$.
 β' = compressibility of bulk phase $[\partial \log V / \partial P]$.

(1) As figs. 3, 4 and 5 show, a linear relation between $\log \phi$ and a is well maintained at all temperatures. Table II shows the value of $\log \phi_0$ obtained by extrapolating the separate lines to $a = 0$.

TABLE II

	25°	30°	50°	70°
Benzene	3.476	—	3.475	3.475
Acetone	—	3.824	—	3.820
Methyl alcohol	3.690	—	3.680	3.660

ϕ_0 , which is of course independent of the evaluation of the surface area, is seen from this table to be also independent of temperature for benzene and acetone. A linear relationship between $\log \phi$ and a does not in fact extend to $a = 0$ for the powdered adsorbent (Palmer and Clark 1935, p. 377): an upward trend near $a = 0$ is already noticeable in the benzene diagram (fig. 3). A previous suggestion (Palmer 1937, p. 263) that this divergence is attributable to surface inhomogeneity and would probably not be shown by an ideal plane surface is supported by this new evidence that ϕ_0 is independent of T , which can hardly be accidental. On the other hand, for methyl alcohol the apparent change of ϕ_0 with T is well outside experimental error. This is regarded as abnormal behaviour, and is explained by association in the film (p. 200 below). Fig. 2, which shows the results of experimental method (b) applied to acetone, indicates that at least for this substance the value of $-\partial \phi / \partial T$ increases somewhat with rise of temperature, but the inevitable experimental error renders a quantitative estimate of the variation uncertain. It may be assumed that for films on silica, where no association or other chemical change occurs in the film,

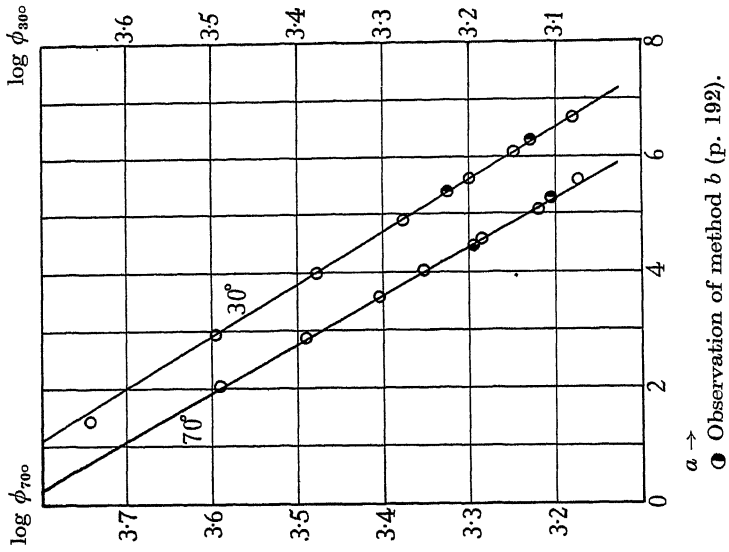


Fig. 4. Acetone.

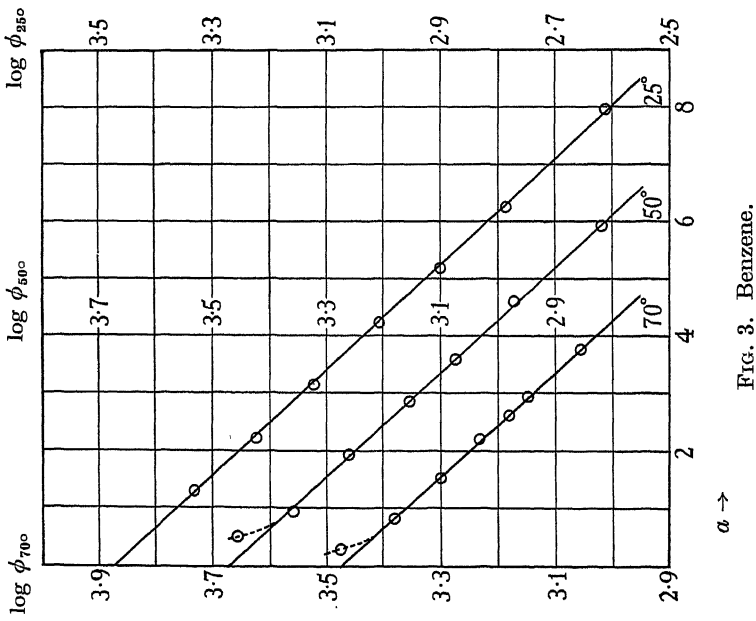


Fig. 3. Benzene.

● Observation of method b (p. 192).

the isothermals can be represented by straight $\log \phi$, a lines radiating fan-wise from a common origin ϕ_0 , the slope of the lines increasing with increase of temperature (i.e. $[\partial \phi / \partial T]_a$ negative).

(2) *The thermal expansion of the film.* We define the thermal expansion α_ϕ by

$$\alpha_\phi = [\partial \log s / \partial T]_\phi = -[\partial \log a / \partial T]_\phi = [\partial \log A / \partial T]_\phi.$$

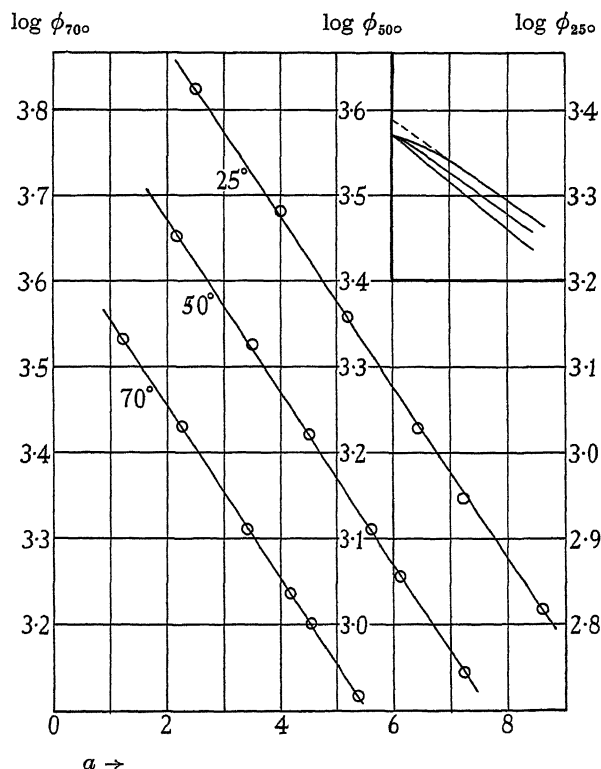


FIG. 5. Methyl alcohol.

TABLE III

	Slope of $\log \phi$ line at				Mean α	α' (liquid, at normal pressure)
	25°	30°	50°	70°		
Benzene	0.1058	—	0.1079	0.1100	0.00090	0.00117
Acetone	—	0.1115	—	0.1185	0.00152	0.00132
Methyl alcohol	0.1080	—	—	—	—	—

For benzene and acetone α is independent of a , but, as fig. 2 suggests, probably rises with increase of temperature. The values of α agree in mag-

nitude with those found by Polanyi and Goldmann (1928) and by Coolidge* in their work on charcoal adsorption of vapours, but their evidence on the dependence of α on a and T is conflicting. Thus Coolidge's data give $\alpha_{100^\circ}^{50^\circ}$ nearly independent of a for benzene and CCl_4 , and $\alpha_{60^\circ}^{50^\circ}$ independent of a for $(\text{C}_2\text{H}_5)_2\text{O}$. Neither of these authors investigated acetone; for adsorbents other than those mentioned their results fluctuate irregularly. Coolidge's data give $\alpha_{150^\circ}^{60^\circ} = 0.0013$ for benzene for a range of ϕ from 6000 to 3000 cal.

In attempting to elucidate the physical state of the film by comparison of its properties with those of liquids and compressed gases, account must be taken of the effect of orientation of permanent or induced dipoles in the surface field. We may regard the potential ϕ as consisting of three terms, $\phi = \phi_0 - E_i - E_T$, where E_i is the interaction energy due to orientation, and E_T the thermal interaction energy. If we imagine a film in equilibrium with a vapour phase at pressure p and confined on the surface by a lateral barrier which exerts on the film a force F , the barrier may be moved inward isothermally while p is simultaneously increased so that no vapour molecules leave or enter the film. The work done on the film is $\int F dA = \Delta E_i + \Delta E_T$.

Now if the adsorbate is non-polar but has isotropic polarizability ΔE_i will not depend on temperature, since the induced dipoles will maintain positions of minimum potential energy in spite of thermal agitation; with a polar substance, however, thermal disturbance must always tend to give fluctuations between positions of potential energy greater than the minimum, and ΔE_i must increase with rise of temperature. It follows that $(\partial\phi/\partial T)_a$ and therefore $\partial \log s / \partial T$ must be relatively larger for a polar than for a non-polar adsorbate. For a bulk liquid or compressed gas where no surface field imposes orientation these special effects would not exist. In agreement with these views

$$\alpha_{\text{acetone}} - \alpha_{\text{benzene}} > \alpha'_{\text{acetone}} - \alpha'_{\text{benzene}}.$$

A striking example of the marked effect of dipole moment upon thermal expansion in a film is furnished by methyl alcohol where in contrast with benzene and acetone the film shows at first a large expansion that, however, decreases rapidly with increasing ϕ :

Thermal expansion of methyl alcohol films (50–70°)

ϕ (cals.)	α	
3000	0.0022	
2000	0.0013	[$\alpha' = 0.00113$]
1000	0.0010	

* Polanyi and Goldmann (1928) calculate and quote values of α obtained from the data of Coolidge, who did not himself make the calculations.

Later discussion (p. 203 below) will indicate that molecular association in this film at low values of a greatly strengthens the normal (monomeric) dipole, a behaviour reflected in the high initial value of α . Coolidge's data also give $\alpha_{90}^\circ = 0.0023$ for $\phi = 3000$ for the alcohol on charcoal.

There is certainly a general agreement in the order of magnitude of α and α' but it is doubtful whether for this reason alone the film can be considered to be in the liquid state (cf. Polanyi and Goldmann 1928). While there exists no rational method of correlating F with the hydrostatic pressure P in dynes cm.², it is generally agreed that P is of the order 100–1000 atm. and must increase rapidly with a . Now data for bulk liquids show that α' decreases markedly with increase in pressure, but is little influenced by temperature, while for films of benzene or acetone the above data suggest an exactly contrary behaviour.

From the integrated expression for F previously given (Palmer 1937) it is easily seen that

$$(FA)_\phi = g[(\phi_0 - \phi)/c - \phi] = \text{constant, independent of } T,$$

where g converts to suitable units, and $c = \log \phi_0/\phi$. If therefore points are chosen on isothermals such that $\phi = RT \log p_s/p$ has a constant value a simple Traube type of expression connects F with A at such points. Hence $-(\partial \log F/\partial T)_\phi = \partial \log A/\partial T$, and F_ϕ decreases with rise of temperature.

(3) *Compressibility (β) and value of $[d \log F/\partial T]_a$.* A comparison of $\partial \log F/\partial T$ as a function of F with $\partial \log P/\partial T$ as a function of P in the bulk phase avoids the difficulty of relating P to F ; with the same advantage we may compare β and β' as functions of these variables respectively. For a perfect gas $\partial \log P/\partial T = \frac{1}{T} \sim 0.003$, while calculation from Pv data for CO₂ under high pressure* shows that under such conditions this coefficient is much greater, but decreases with increasing pressure, being, at 25°, 0.035 for 170 atm. and 0.014 for 860 atm. Table IV gives values for an acetone film at 30°.

TABLE IV

θ	F	$\partial \log F/\partial T$
0.2	5	0.00235
0.4	18	0.00122
0.6	35	0.00077
0.8	51	0.00043

* *International Crit. Tables*, 3, 11.

The values of the coefficient tend at low θ towards that for a perfect gas and are all much less than those for compressed CO_2 , but they show, like CO_2 , a marked fall with increasing F .

For a gas under low pressure $\beta' = -1/p$, and for CO_2 between 200 and 1000 atm. $-\beta' = 0.4/p$;^{*} for the gaseous state we may thus assume an approximately linear relation between $-\beta'$ and p^{-1} . For acetone in bulk $-\beta' = \text{const.}/p^{0.32}$ for P between 250 and 1500 atm.; for benzene the very similar relation $-\beta' = \text{const.}/p^{0.30}$ holds for P between 200 and 600 atm.[†] For both liquids the index to p lessens at lower pressures. When $\log \beta$ is plotted against $\log F$ for benzene films we find a slope of unity up to $A = 50$ and $F = 7$ dynes, and then a rather sudden change of slope to 0.55, which remains nearly constant to $(A) = 15$ and $F = 40$ dynes; for acetone films the result is the same, except that the change of slope occurs at 18 dynes, but this is not surprising as acetone develops a higher force than benzene. We have thus for the films $-\beta = \text{const.}/p^n$, where $n = 1$ for $\theta < 0.5$ and $n = 0.55$ for greater θ . The state of the film revealed by these data is more nearly akin to that of a compressed gas than that of a liquid.

(4) *Heats of adsorption.* (a) *Benzene and acetone.* For these adsorbates, $\log_{10} \phi = \log_{10} \phi_0 - sa = \text{const.} - sa$

$$-T(\partial\phi/\partial T)_a = T\phi a(\partial s/\partial T) \times 2.306$$

$$L_a = L + \phi - T(\partial\phi/\partial T)_a,$$

where L = the molar latent heat of evaporation of the bulk liquid at T° . $(dH/da)_a = L_a$ = the heat liberated when 1 g. mol. of vapour is adsorbed from a large volume of vapour in equilibrium with a film of indefinitely large extent and already containing a micromols. adsorbate per sq. m.

$L_a - L = \phi - T(\partial\phi/\partial T)_a$ has been termed the heat of wetting (Polanyi) or the nett heat of adsorption (Coolidge). For the three substances concerned in the present work $\partial\phi/\partial T$ is negative at all values of a , and therefore $L_a - L$ is greater than ϕ . In calculating, a mean value of $\partial\phi/\partial T$ for the temperature range has been used.

(b) *Methyl alcohol.* For this substance the slopes of the $\log \phi$, a lines are very nearly the same at all temperatures, and the principal consideration is $\partial\phi_0/\partial T$. As mentioned above the apparent ϕ_0 decreases with rise of temperature (Table II). The reason for this is suggested by earlier work (Palmer and Clark 1935; Palmer 1937, p. 256), where it is shown that in

* *International Crit. Tables*, 3, 11.

† *International Crit. Tables*, 3, 36 and 42 resp.

the first part of the adsorption of methyl and ethyl alcohols $\log \phi$ decreases more slowly than in the ultimate linear relation with α , an effect then attributed to association in the film, and one which would be expected to diminish in importance rapidly with rise of temperature. We may conclude that the true ϕ_0 is near to the value ($\log \phi_0 = 3.660$) given by extrapolating the $\log \phi$ line for 70° , and that the lines for 50 and 25° undergo increasing curvature towards this common value as α tends to zero (represented diagrammatically in the inset to fig. 5). In support of this view the heat of

TABLE V. HEATS OF ADSORPTION

ϕ	$T(\partial\phi/\partial T)$	$L_a - L$	L_a	θ
BENZENE. $L = 8050$ cal.				
2990	0	2990	11,140	0.0
2512	130	2642	10,692	0.1
2120	237	2357	10,407	0.2
1784	276	2060	10,110	0.3
1503	310	1813	9,863	0.4
1267	326	1593	9,643	0.5
1068	330	1398	9,448	0.6
898	324	1222	9,272	0.7
757	312	1069	9,119	0.8
637	296	933	8,983	0.9
537	277	814	8,864	1.0
ACETONE. $L = 7710$ cal.				
6640	0	6640	14,350	0.0
5540	465	6005	13,715	0.1
4620	775	5395	13,105	0.2
3859	949	4808	12,518	0.3
3220	1080	4300	12,010	0.4
2688	1126	3814	11,524	0.5
2240	1130	3370	11,080	0.6
1873	1099	2972	10,682	0.7
1560	1050	2610	10,320	0.8
1305	986	2291	10,000	0.9
1089	913	2002	9,712	1.0
METHYL ALCOHOL. $L = 8946$ cal.				
4570	0	4570	13,516	0.0
3470	2100	5570	14,516	0.2
2630	2190	4820	13,766	0.4
2239	1862	4101	13,047	0.5
1910	1580	3490	12,436	0.6
1622	1349	2971	11,917	0.7
1380	1150	2530	11,576	0.8
1175	977	2152	11,098	0.9
1000	832	1832	10,778	1.0

adsorption shows an initial rise, which has been shown theoretically to indicate negative interaction energy in the film. For the purposes of calculation the initial part of the $25^\circ \log \phi$ line is taken as given in form by the earlier work, and the 70° line as linear right up to $a = 0$. The values of the heat of adsorption in the neighbourhood of its maximum are thus only approximate, but there appears to be no doubt of the existence of the maximum.

The heats of adsorption are somewhat lower and more differentiated according to the type of substance than those found with vapours and charcoal: Pearce and McKinley (1928) find by direct calorimetry at low adsorptions the approximate values, benzene 17,000 cal., acetone 16,000 cal., methyl alcohol 16,800 cal.

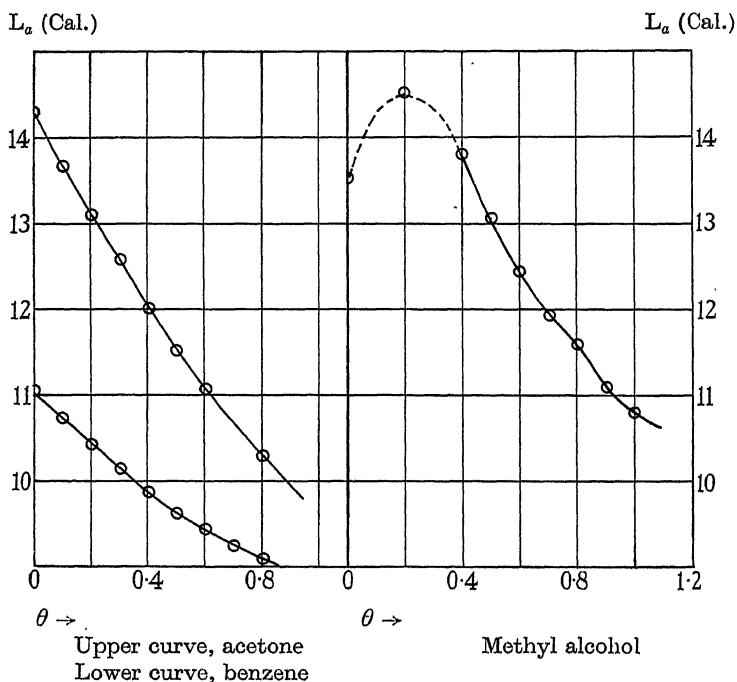


FIG. 6

It will be seen from fig. 6 that the heat for benzene and acetone decreases in nearly linear relation to θ up to $\theta \sim 0.5$: the slight divergence from a straight line is within the limits of error in determining $T\partial\phi/\partial T$. Such a relation between L and θ is in accordance with the recent calculations of J. S. Wang (1938) when the interaction energy is small (less than $kT/4$), but it may be doubted whether the vitreous silica surface presents the orderly array of sites for adsorption that Wang's calculation presumes.

In our case, where more than one layer can be built up, it seems very probable that at $\theta \sim 0.5$ a second layer begins to be formed owing to the increasing difficulty of entering the first layer against the interaction forces. Ultimately at higher pressures the first layer will gradually fill. It is significant that at $\theta \sim 0.5$ the relation of the compressibility β to F changes rather sharply (p. 200 above) and also the form of the equation of state for the film alters (Palmer 1937).

In connexion with the heat curve for methyl alcohol it is interesting to note some recent measurements of the molecular polarization of dissolved alcohols (Wolf and Herold 1934; Hennings 1935; Romanov and Eltzin 1937) where it is proved that the process of association may lead to a large increase of dipole moment, which is especially marked with methyl alcohol. As the potential ϕ is thought to depend on μ^2 (Palmer 1937, p. 265), if the moment of a dimeric complex is greater than that of a simple molecule by 50 % the potential for the complex will be greater than the sum for two simple molecules, and the curvature of the $\log \phi$ line is explained. The maximum value in the polarization curves for the dissolved alcohols decreases more rapidly than in proportion to $1/kT$, which is in agreement with the rapid loss of curvature in the $\log \phi$ lines with rise of temperature.

There do not appear to be any comparable observations for methyl alcohol on other adsorbents. Neither Pearce and McKinley (1928) nor Coolidge (1924) provide data for calculation in the range of adsorption where a maximum might have been expected. Kruyt and Modderman (1930) from the experiments of Ray (1925) on adsorption of N_2O_4 — NO_2 mixtures on silica gel, where association might well be expected, find a marked maximum in the heat when calculating from isothermals at 15 and 57°, but only a steady fall when those at 57 and 80° are used. The data of Wolf and Herold indicate that the polarization falls very rapidly from the maximum on the side of higher concentration, and we may assume a correspondingly rapid fall of mean moment in the adsorption film. Wang (1938) does not consider the case of increasing moment nor of actual chemical association, but shows semi-quantitatively that a rapid decrease of moment would lead to inflexions in the heat curve, such as is actually seen in fig. 6 (methyl alcohol) at $\theta = 0.8$.

(Added 28 July 1938.) The relation of the heat of adsorption dH/da to the amount adsorbed, which in the absence of association may be expressed in the form $d^2H/da^2 = -f(1/a)$, where $f'(a)$ is positive, is in good qualitative agreement with the earlier investigations of Coolidge (1924), and especially with those of Polanyi and Goldmann (1928), on the adsorption of vapours

upon charcoal, although in the latter the complex relation between $\log \phi$ and a rendered extrapolation to $a = 0$ impossible. Bangham and Mosallam, in a recent communication (1938*b*), report that dH/da for benzene and methyl alcohol on surfaces of mica differs little from the normal latent heat when the adsorption has reached greater amounts than correspond approximately to one molecular layer, a result corroborated in Table V above. They also find 12,000 cal./mol. and 13,400 cal./mol. respectively for the initial heats of adsorption of benzene and methyl alcohol, with which our estimates of 11,140 cal./mol. and 13,520 cal./mol. are in excellent agreement, in spite of the different nature of the surfaces, but they claim that these values are maintained nearly constant throughout the formation of the first layer, and then fall abruptly. In the case of benzene their isothermals exhibit even in the first layer most marked discontinuities in slope, on the significance of which they lay some stress. Now it appears to the writer that the evidence so presented is conflicting, as the discontinuities must imply some kind of interaction between the adsorbed molecules, since heterogeneity of surface is not likely to lead to breaks except in the unlikely case of micro-surfaces associated with very widely separated "levels" of adsorption energy, and such interaction is incompatible with a constant heat of adsorption.

In an earlier paper (1938*a*) Bangham and Mosallam show that the benzene and methyl alcohol isothermals conform at low pressures to a simple Langmuir formula, and associate this with the constant heat of adsorption. However, an inspection of the literature reveals that, from an experimental standpoint, a decreasing heat of adsorption is by no means incompatible with a simple Langmuir formula (i.e. with only two constants). Thus Zeise (1928) showed that Titoff's data (1910) for adsorption of CO_2 , NH_3 and N_2 on charcoal (outgassed at 550°) over a wide range of temperature fit simple Langmuir expressions with exactness over the whole range of pressure used by Titoff; from this we may conclude that only unimolecular layers are concerned. McBain and Britton (1930) have fully confirmed the agreement for nitrogen with pressures rising to 60 atm. Titoff's own direct calorimetric measurements, however, showed (1910) that the heats of adsorption of these gases decreased steadily as follows throughout the range of pressure used: CO_2 , 8100–6400 cal./mol.; NH_3 , 11,000–7000 cal./mol. and N_2 , 8000–4000 cal./mol. These changes are all fully confirmed by indirect calculation from Titoff's isothermals (Kruyt and Modderman 1930). Magnus and Kälberer (1927), using charcoal outgassed at 600° , found the heat for CO_2 to fall steadily from 12,500 cal./mol. at very small pressures and adsorptions to 6720 cal./mol. at higher pressures. It seems clear from these and other

similar experimental facts that the Langmuir formula is not actually restricted either to a plane surface or to cases of constant heat of adsorption, which as a matter of fact appear to be rare. The constant c_2 in the formula $a = c_1 c_2 p / (1 + c_2 p)$ depends on the ratio $\alpha / e^{-L_a/RT}$, where α is the "accommodation coefficient" for the condensation process, and the constancy of c_2 means that α must vary rapidly with a (and θ) to compensate the change in L_a .

(*Note added in proof.*) In a note in their paper (1938*b*) Bangham and Mosallam refer to the work of Palmer and Clark (1935), and no doubt by an oversight in which specific surface was confused with actual working surface are led to an erroneous statement about the area of our silica powder. We used *ca.* 15 g. of powder of specific surface *ca.* 5000 cm.², giving a total working surface of *ca.* $15 \times 5000 = 75,000$ cm.², which is almost exactly four times the area of the mica used by Bangham and Mosallam, and not one-quarter of their surface as stated in the note.

SUMMARY

1. Isothermals have been obtained at temperatures between 70 and 25° for adsorption of the vapours of benzene, acetone and methyl alcohol on known surfaces of vitreous silica, by methods giving a check on the constancy of the surface condition.

2. The values of the coefficients of thermal expansion, compressibility and the temperature variation of the lateral force have been calculated and discussed in relation to parallel properties of the bulk phase.

3. The heats of adsorption have been calculated from the temperature variation of the adsorption potential. Their relations to the fraction of surface covered are in agreement with theoretical predictions for films with interaction energy.

REFERENCES

- Bangham and Mosallam 1938*a* *Proc. Roy. Soc. A*, **165**, 552.
 — — 1938*b* *Proc. Roy. Soc. A*, **166**, 558.
 Coolidge 1924 *J. Amer. Chem. Soc.* **46**, 596.
 Hennings 1935 *Z. phys. Chem. B*, **28**, 267.
 Kruyt and Modderman 1930 *Chem. Rev.* **7**, 259.
 McBain and Britton 1930 *J. Amer. Chem. Soc.* **52**, 2217.
 Magnus and Kälberer 1927 *Z. anorg. Chem.* **164**, 345, 357.
 Palmer 1937 *Proc. Roy. Soc. A*, **160**, 254.
 Palmer and Clark 1935 *Proc. Roy. Soc. A*, **149**, 360.

- Pearce and McKinley 1928 *J. Phys. Chem.* **32**, 360.
 Polanyi and Goldmann 1928 *Z. phys. Chem. A*, **132**, 321.
 Ray 1925 *J. Phys. Chem.* **29**, 74.
 Romanov and Eltzin 1937 *Phys. Z. Sowjet.* **11**, 526.
 Titoff 1910 *Z. phys. Chem.* **74**, 641.
 Wang, J. S. 1938 *Proc. Camb. Phil. Soc.* **34**, 238.
 Wolf and Herold 1934 *Z. phys. Chem. B*, **27**, 58.
 Zeise 1928 *Z. phys. Chem. A*, **136**, 408.

The rates of transformation in ethyl alcohol of ammonium and ethylammonium cyanates to the corresponding ureas

BY CHRISTINA C. MILLER AND JAMES R. NICHOLSON

Chemistry Department, University of Edinburgh

(Communicated by J. Kendall, F.R.S.—Received 29 June 1938)

In aqueous and aqueous ethyl alcoholic solutions ammonium and alkylammonium cyanates seem to be converted to the corresponding ureas by a reversible bimolecular mechanism associated primarily with univalent ammonium (or alkylammonium) and cyanate ions. Thus, if C_A and C_B are the concentrations of the ionic reactants A and B , based on the conception of complete dissociation, and the reverse reaction is negligible, the velocity v of the forward reaction is represented by the equation, $v = k_0 C_A C_B F$, where k_0 is the velocity constant, and F is a reaction kinetic factor which is at least qualitatively represented by $f_A f_B / f_X$, as defined by Brönsted (1922). f_A , f_B and f_X represent, respectively, the activity coefficients of the reactants and a critical complex X , which, in our case, is a neutral molecule of which the activity coefficient may be taken to be unity. For example, when the logarithm of k_C , the so-called velocity constant determined experimentally at various concentrations from the equation $v = k_C C_A C_B$ ($k_C = k_0 F$), is plotted against the square root of the average ionic strength (here, simply the average concentration) of the interval to which it refers, a curve is obtained which at low concentrations approximates to the linear relationship required by the Debye-Hückel theory. Miller (1934, 1935) showed this for methylammonium cyanate in water and in 98.1 % aqueous

ethyl alcohol. Warner and Stitt (1933) obtained a similar result for ammonium cyanate in water. The rates of transformation of ammonium and methylammonium cyanates in water were influenced by neutral salts in the manner predicted by Brönsted.

A peculiarity was observed in Miller's experiments with methylammonium cyanate in alcoholic solution, which were made over the concentration range 0.0002–0.05 M at two temperatures. k_0 was determined by graphical extrapolation and F , as determined from the equation $k_G = k_0 F$, was then compared with the value of f^2 , calculated for a univalent ion from the equation of Gronwall, La Mer and Sandved (1928). It was assumed that f_x could be considered throughout to be unity. At 0° and at 10° F equalled f^2 when reasonable values of a , the mean distance of closest approach of other ions to that under consideration, were taken, but the values at the two temperatures were 6.1 and 7.6 Å, respectively. A further unexpected result was that at infinite dilution Christiansen's (1924) equation, omitting the steric and orientation factors, for a bimolecular reaction of measurable velocity between univalent ions of opposite sign satisfactorily represented the experimental results only when the distances, r , between the centres of colliding and reacting ions were set equal to a .

In view of the above results it was considered advisable to extend the experiments to other cyanates and to have data for at least three temperatures. Accordingly further experiments were undertaken for ammonium and ethylammonium cyanates in the same solvent as was used for methylammonium cyanate, namely aqueous ethyl alcohol containing 98.1 % by weight of ethyl alcohol.

EXPERIMENTAL

The experimental arrangements were essentially the same as before. Two preparations of the solvent were required, but all the experiments for each cyanate were made with a single preparation of conductivity 0.1 gemmho at 10°. Ammonium cyanate was prepared from pure ammonium chloride and cyanuric acid and could be kept for several days. Ethylammonium cyanate was prepared from cyanuric acid and ethylammonium chloride which had been freed from ammonium and other alkylammonium chlorides. Each velocity experiment required a fresh preparation.

The methods of determining the rates of reaction were exactly as previously described, the concentration range 0.05–0.003 M being covered by means of chemical analysis, and the range below 0.001 M by a conductivity method. As the equilibrium point concentrations of the reactions of both

cyanates were of the same order as that for methylammonium cyanate, the reverse transformations of the ureas into cyanates could be ignored.

For ammonium cyanate the three temperatures chosen were 16, 24 and 32° C, and for ethylammonium cyanate, 0, 10 and 20° C. The equivalent conductivity of the cyanates was determined at 16 and 32° C for ammonium cyanate, and at 0 and 20° C for ethylammonium cyanate. For 24 and 10° C interpolated values were used. Temperature was controlled to $\pm 0.01^\circ$ C and recorded on thermometers that had been compared with N.P.L. standard ones.

TABLE I
AMMONIUM CYANATE

16.00° C		24.00° C		32.00° C	
k_C	C_{average}	k_C	C_{average}	k_C	C_{average}
0.000640	0.04161	0.001936	0.03441	0.00530	0.03209
0.000728	0.03027	0.002102	0.02669	0.00591	0.02449
0.000904	0.01766	0.002607	0.01784	0.00564	0.02340
0.001066	0.01049	0.00319	0.00944	0.00672	0.01699
0.001683	0.002829	0.00495	0.002836	0.00817	0.00918
0.00216	0.000940	0.00566	0.001035	0.01130	0.003059
0.00224	0.000746	0.00653	0.000440	0.01291	0.000924
0.00244	0.000485	0.00754	0.000191	0.01534	0.000481
0.00286	0.000241	0.00865	0	0.0197	0.000220
0.00294	0.000183			0.0202	0
0.00309	0				
				*0.0218	0.000182
				*0.0235	0.000154

ETHYLAMMONIUM CYANATE

0.00° C		10.00° C		20.00° C	
k_C	C_{average}	k_C	C_{average}	k_C	C_{average}
0.000672	0.03795	0.002353	0.03408	0.00769	0.03233
0.000775	0.02748	0.002673	0.02603	0.00776	0.02958
0.000916	0.01868	0.003015	0.01795	0.00827	0.02504
0.001095	0.01138	0.00371	0.00911	0.00891	0.01776
0.001674	0.003278	0.00515	0.002697	0.00916	0.01696
0.00193	0.000999	0.00636	0.000595	0.01025	0.01001
0.00219	0.000517	0.00718	0.000275	0.01346	0.002759
0.00213	0.000211	0.00624	0.000124	0.01525	0.000895
0.00216	0.000220	0.00847	0	0.01577	0.000479
0.00280	0			0.0132	0.000208
				0.0182	0.000182
				0.0215	0

RESULTS

In Table I the data for the velocity experiments are presented. k_G is given as formerly in moles per litre per second and C_{average} is the average concentration of the interval to which it refers.

The equivalent conductivities at infinite dilution of ammonium and ethylammonium cyanates were found by extrapolation from the experimental data to be approximately as follows:

Ammonium cyanate: 32.5 (16°), 36.9 (24°), 41.4 (32°);

Ethylammonium cyanate: 25.8 (0°), 31.1 (10°), 37.8 (20°).

THE MECHANISM OF THE REACTION

In figs. 1 and 2 are plots of $\log k_G$ against $\sqrt{C_{\text{average}}}$ for each cyanate at the three temperatures. The broken lines through the points will be referred to presently.

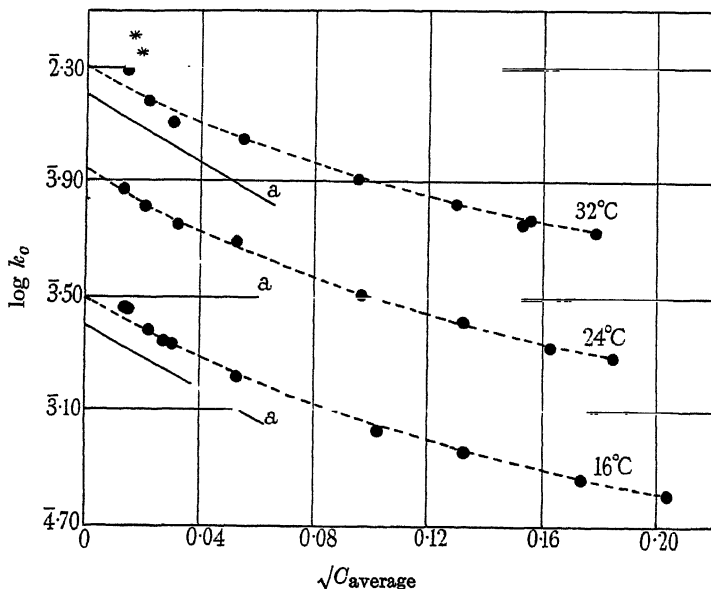


FIG. 1. Ammonium cyanate. *a* Theoretical limiting slope.

In a general way the results resemble those found for methylammonium cyanate, but with ammonium cyanate below a concentration of 0.0004M there is a definite tendency for the rate to increase more than anticipated.

Either the production of urea is being catalysed or another reaction is occurring simultaneously. Traces of impurity in the solvent might be responsible, but experiments at low concentrations repeated at 32° with two fresh preparations of alcohol of the same composition gave similar results. These are shown marked with an asterisk in Table I and in fig. 1. The neglect of an appreciable reverse reaction, or of possible carbonate formation in these conductivity experiments, would affect k_C in the opposite sense to that observed. Owing to the difficulty of experimenting at high dilutions we propose meantime to neglect the data obtained at molar concentrations below about 0.0004. k_0 was determined by graphical extrapolation of smoothed curves drawn through the appropriate points in figs. 1 and 2. We have omitted these curves from the diagrams, but the values deduced for k_0 are given in Table I.

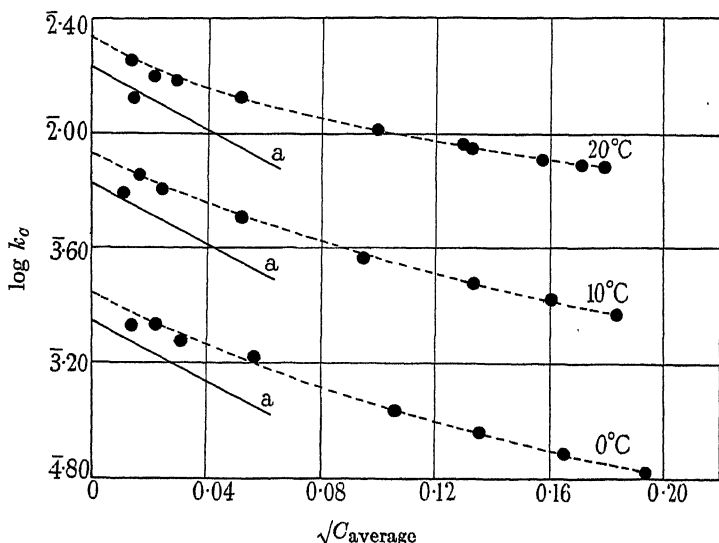


FIG. 2. Ethylammonium cyanate. *a* Theoretical limiting slope.

The kinetic factor F and the mean distance of closest approach of the reactant ions

If $k_C = k_0 F = k_0 f_A f_B / f_X$, $f_A = f_B = f$ for a univalent ion according to the Debye-Hückel theory, and $f_X = 1$, then $\log k_C$ when plotted against $\sqrt{C_{\text{average}}}$ ought to tend at low concentrations towards a straight line of a certain limiting slope. This is shown in figs. 1 and 2. At higher concentrations F may be compared with the values of f^2 derived from Gronwall, La Mer and Sandved's equation. Since it is uncertain what value should be

assigned to the term a (see p. 207) we tested the theory as follows. Using the experimental data in Table I for concentrations above 0.0004M we calculated for various trial values of a the values of f^2 corresponding to C_{average} . That value of a , which for a particular cyanate at one temperature, led to the most nearly constant value of k_0 in the equation $k_C = k_0 f^2$ was accepted as being correct. The values deduced for a in Angström units, and for k_0 are:

Ammonium cyanate: $a = 5.4 \text{ \AA}$, $k_0 = 0.00311$ (16°); $a = 6.0 \text{ \AA}$,
 $k_0 = 0.0086$ (24°); $a = 7.4 \text{ \AA}$, $k_0 = 0.0207$ (32°);

Ethylammonium cyanate: $a = 5.6 \text{ \AA}$, $k_0 = 0.00291$ (0°); $a = 7.1 \text{ \AA}$,
 $k_0 = 0.0088$ (10°); $a = 11.1 \text{ \AA}$, $k_0 = 0.0224$ (20°).

In figs. 1 and 2 the broken lines represent the theoretical relationship between $\log k_C$ and $\sqrt{C_{\text{average}}}$ based on the above values of k_0 and a . It is evident that ammonium and ethylammonium cyanates, like methylammonium cyanate, require values of a with high positive temperature coefficients. With the exception of the value at 20° for ethylammonium cyanate their magnitudes are of an acceptable order, and compare very favourably with the sums of the radii of the appropriate ions calculated from the conductivity data on p. 209 in conjunction with Stokes's law. For ammonium and ethylammonium cyanates at 25° these are respectively 7.5 and 6.8 Å, and they are not subject to a large temperature coefficient.

We have examined the literature for evidence of the temperature variation of a in non-aqueous solvents. From electromotive force data for hydrochloric acid up to a concentration of 0.02M in 45 % aqueous dioxane Harned (1938) derived values of a that decrease $17\frac{1}{2}\%$ for a 50° rise in temperature; in 70 % aqueous dioxane the corresponding decrease is less than 10 %. To such variations no significance was attached, and it would appear that our kinetically deduced values are unusual. A weak point in our deductions is that we have had to depend more on the higher than on the lower concentration data to which the Gronwall equation would normally better apply.

Christiansen's equation at infinite dilution

Christiansen's theory was applied to the experimental results in the same manner as for methylammonium cyanate (Miller 1935), the results for which are given for comparison in Table II. The fourth and seventh columns of the table show that there is no general agreement with the theory whether r is set equal to a or taken to be constant. E the energy of activation in

kilocalories is so variable with temperature that it cannot be a true energy of activation.

TABLE II

Temp. ° C	r in Å	E	$\frac{k_0 \text{ calc.}}{k_0 \text{ exp.}}$	r in Å	E	$\frac{k_0 \text{ calc.}}{k_0 \text{ exp.}}$
Ammonium cyanate						
16	7.5	20.34	2	5.4	24.15	0.004
24	7.5			6.0		
32	7.5	17.27	330	7.4	23.61	0.01
Ethylammonium cyanate						
0	6.8	15.55	1200	5.6	20.46	0.07
10	6.8			7.1		
20	6.8	13.73	30,000	11.1	17.48	40
Methylammonium cyanate						
0.04	6.2	14.87	3000	6.1	18.76	2
9.98	6.2			7.6		

Finally E was deduced at the lowest temperature for each cyanate by substituting for r in the appropriate equation the corresponding value of a . Assuming E to be independent of temperature we then found the values of r required to give agreement between experiment and theory at the higher temperatures to be as follows:

Ammonium cyanate: (5.4), 5.6, 6.2 Å;

Ethylammonium cyanate: (5.6), 6.8 Å, indeterminate.

Hence r requires a positive temperature coefficient distinctly lower than that found for a .

The factors E_A and A of the Arrhenius equation

The velocity constants of a great many reactions in solution may be represented by Arrhenius's equation, $k = Ae^{-E_A/RT}$ where E_A , the critical increment, and A are constants. A is commonly factorized into PZ , Z in a bimolecular reaction being the number of encounters between the reactants, as given by the kinetic theory of gases, and P , a probability factor which may range in value from 10^{-8} to 10^5 . So-called "normal" reactions are associated with probability factors of about unity.

The data in Table III have been derived by inserting in Arrhenius's equation values of k_C obtained from the smoothed curves referred to on p. 210. Miller's data for methylammonium cyanate in the same solvent are

included. For each cyanate is given the theoretical value of $\log Z$ referred to a collision diameter based on conductivity data (p. 211), and the lowest experimental temperature; it does not vary much with moderate changes in α and T . E_A is given in kilocalories.

TABLE III

Cyanate	Temp. range	$C = 0.0004$	0.005	0.015	0.03
Ammonium $\log Z = 11.9$	16–24°	$E_A = 21.6$	22.7	22.7	22.0
		$\log A = 13.7$	14.3	14.1	13.5
	24–32°	$E_A = 19.2$	19.6	21.3	22.0
		$\log A = 11.9$	12.1	13.1	13.5
Ethylammonium $\log Z = 11.7$	0–10°	$E_A = 17.2$	17.2	18.0	18.6
		$\log A = 11.1$	10.9	11.4	11.8
	10–20°	$E_A = 15.2$	16.1	17.5	18.7
		$\log A = 9.6$	10.1	11.0	11.8
Methylammonium $\log Z = 11.6$	0–10°	$E_A = 16.3$	17.0	17.2	17.6
		$\log A = 10.7$	11.0	11.1	11.2

Considering the data as a whole there are no highly abnormal values of E_A and $\log A$, showing that the reactions are probably not complex. At the highest concentration only, for ammonium and ethylammonium cyanates, is the Arrhenius equation satisfied. Except for ammonium cyanate between 16 and 24° E_A and $\log A$ increase throughout with increase in concentration, the rate of increase increasing as temperature rises. As E_A changes with concentration, with temperature, and from one cyanate to another, $\log A$ varies in the same direction, but not always to the same extent. Such parallel variations are by no means uncommon. The maximum deviation of $\log A$ from $\log Z$, approximately ± 2 , is not excessive. It has been shown by Moelwyn-Hughes (1936) that in a bimolecular reaction involving ions P will differ from unity, being greater at ordinary temperatures and low concentration when the reactants are oppositely charged. His formula for two univalent ions at a concentration of 0.0004M gives for 98.1% aqueous ethyl alcohol at 10°, and a collision diameter of 6.8 Å, $\log P = 1.77$. Only the result for ammonium cyanate between 16 and 24° satisfies this condition.

The variation of E_A and $\log A$ with concentration

According to Moelwyn-Hughes's theory cited above, which is based on the fulfilment of the Brönsted-Debye-Hückel theory, E_A and $\log A$ should decrease as concentration increases, when the ionic reactants are of opposite sign, and increase if they are of the same sign. Our results show almost

entirely the reverse of the predicted effect. Further, the divergences from the theory increase as temperature rises.

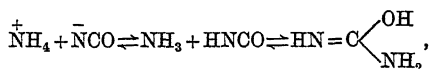
For ammonium cyanate in water and in 28.3 % aqueous methyl alcohol Svirbely and Warner (1935) and Svirbely and Schramm (1938) show values of E_A that change in the predicted manner, when $C = 0$ and 0.0376. This is, however, bound to occur, because the values of k_C and k_0 from which the corresponding values of E_A are deduced are related by means of the formula, $k_C = k_0/1 + 2A\sqrt{C}$, $1/1 + 2A\sqrt{C}$ being Scatchard's (1930) formula for $f_A f_B / f_X$, based on the Debye-Hückel theory. We have very carefully examined a few unpublished experiments of Miller for methylammonium cyanate in water at 50°, and we consider that, there too, E_A tends to increase slightly with increasing concentration (0.05–0.08M), just as it does in ethyl alcohol. Consequently we are of the opinion that the effect we have noted is real, and associated with the non-fulfilment of the Brönsted-Debye-Hückel theory (α constant).

The variation of E_A and $\log A$ with temperature

La Mer (1933) pointed out that Tolman's statistical development demands in general that E_A must be a function of temperature. Our results for ammonium and ethylammonium cyanates show for E_A and $\log A$ large negative temperature coefficients at low concentrations and none at the highest—0.03 M. At still higher concentrations it would appear that temperature coefficients might become positive. In the literature there is a lack of kinetic data, suitable for comparison purposes, for ionic reactions in solvents like alcohol at several temperatures and ionic strengths, but the following may be of interest. v. Kiss and Bossányi (1935) show for a reaction between two negative ions in 30 % aqueous ethyl alcohol at one ionic strength (0.005) a small negative temperature coefficient for E_A . For the same reaction in aqueous solution at an ionic strength of 0.02 La Mer and Kamner (1935) find negative temperature coefficients for E_A (9 cal./°C) and $\log A$. Percival and La Mer (1936) found in a catalytic reaction in aqueous solution at 0–40°, involving a positive and a negative ion, that, at an ionic strength of 0.2, E_A and $\log A$ have positive temperature coefficients ($dE_A/dT = 63$ cal.).

Our temperature coefficients for low concentrations are very high compared with those cited. La Mer and Miller (1935), having surveyed the literature, state "that reactions which have so far demonstrated a marked temperature dependence of E_A fall roughly into two classes; namely, those involving a halogen compound and those which are either a hydrolysis or a proton change". Lowry (1934) has suggested that the migration of a proton

from an ammonium ion to a cyanate ion, forming ammonia and cyanic acid, and then urea,



may explain the transformation of ammonium cyanate. Walker and Hambly (1895) found, however, that added ammonia had very little influence on the transformation.

*The relative rates of transformation of the cyanates and
the solvent effect*

Owing to the temperature variation of E_A and $\log A$ a discussion of the relative rates of reaction of ammonium, methylammonium and ethylammonium cyanates, bringing in these factors, cannot be entirely satisfactory, more especially since all the data cannot be referred to one temperature range. Nevertheless some interesting facts emerge in considering Table IV wherein we have collected, in addition to our latest results, all the relevant published data (as far as we know). E_A in kilocalories and $\log A$ refer mainly to zero concentration and to temperature ranges as close as possible to the temperature associated with k_0 .

TABLE IV

Solvent	Cyanate	k_0	E_A	$\log A$	Temp. interval to which E_A and $\log A$ refer
Water	Ammonium	0.000130 (32°)*	23.6	13.0	30–70°
		0.000362 (40°)*†			
		0.0024 (60°)†			
	Methylammonium	0.00046 (40°)‡	22.6	12.4	40–50°
28.3 % CH ₃ OH + 71.7 % H ₂ O	Ethylammonium	0.0022 (60°)†	22.4	12.6	30–60°
		0.0014 (60°)†			
		0.00032 (32°)§			
	Ammonium	0.00128 (10°)	22.0	14.1	16–24°
98.1 % C ₂ H ₅ OH + 1.9 % H ₂ O	Ammonium	0.0202 (32°)	19.1	12.0	24–32°
		0.0154 (10°)¶	16.4	10.8	0–10°
	Ethylammonium	0.0085 (10°)	17.0	11.0	0–10°

* Svirbely and Warner (1935). Result at 32° interpolated.

† Walker and Appleyard (1896). Values very approximate.

‡ Miller (1934). E_A and $\log A$ refer to a concentration of 0.082M.

§ Svirbely and Schramm (1938). Result interpolated.

|| By extrapolation from the data at 32–16° in this paper.

¶ Miller (1935).

The three cyanates react at similar rates in water (k_0 at 40 and 60°). Only for ammonium and methylammonium cyanates can we compare E_A and $\log A$, both of which are slightly greater for the ammonium salt. When the solvent is changed to ethyl alcohol, in which the rates of all three cyanates are increased, and the rates compared at 10°, the most striking feature is that ammonium cyanate reacts about ten times more slowly than the other two cyanates, which bear to each other much the same relationship as they do in water (60°). E_A and $\log A$ are markedly less for methyl- and ethylammonium cyanates than they are for ammonium cyanate. A comparison of rates at one temperature (32°) in three solvents is possible only for ammonium cyanate; as k_0 increases $\log A$ and E_A concomitantly decrease. According to Ogg and Rice (1937), if the same reaction may take place either with solvated or with unsolvated reactants then E_A and A will be greater and k less for the former. In reactions such as ours, where polar reactants form much less polar activated complexes and final products, extensive solvation of the reactants may be expected to occur, and to a different extent in different solvents. The changes in k_0 , E_A and A for ammonium cyanate in water and ethyl alcohol may then be explained on the basis that the degree of solvation is less in alcohol than in water. Now the sizes of the ionic reactants as deduced from conductivity data and Stokes's law are much greater for alcoholic than for aqueous solutions, which might be taken to indicate a higher degree of solvation in alcohol. There is, however, considerable doubt as to the meaning of the large sizes so calculated for non-aqueous solvents (Davies 1933). We note that Verhoeck (1934) studying a reaction in water and ethyl alcohol involving an ionic reactant, explained similar changes in E_A to ours by assuming a higher degree of solvation in water. Since the heat of solvation is usually positive one would deduce for E_A a negative temperature coefficient. For ammonium and ethylammonium cyanates in ethyl alcohol this has already been shown for low concentrations (p. 214).

As cyanate ions are common to the three reactions under consideration it seems justifiable to associate the differences in rates with the cations, the mobilities of which in water and 98 % aqueous ethyl alcohol at 25° are approximately as follows:

Water: NH_4^+ , 75; CH_3NH_3^+ , 62; $\text{C}_2\text{H}_5\text{NH}_3^+$, 50;

Alcohol: NH_4^+ , 19; CH_3NH_3^+ , 27; $\text{C}_2\text{H}_5\text{NH}_3^+$, 23.

It is at once noted that NH_4^+ undergoes a relatively much greater decrease in mobility than the others in the change from water to alcohol. If it is permissible to assume some sort of relationship between ionic mobilities

and the extent of solvation then, if the three ions are similarly solvated in water, presumably NH_4^+ is relatively more solvated than CH_3NH_3^+ and $\text{C}_2\text{H}_5\text{NH}_3^+$ in alcohol. Since k_0 , E_A and A for ammonium and methylammonium cyanates are similar in water one might therefore expect to find the rate for ammonium cyanate in alcohol less than that for methylammonium cyanate, and E_A and A considerably greater. This is actually observed. As k_0 , E_A and A for methylammonium and ethylammonium cyanates are similar in alcohol, and the relative changes in mobility of CH_3NH_3^+ and $\text{C}_2\text{H}_5\text{NH}_3^+$ in passing from alcohol to water are similar, one would predict for methylammonium and ethylammonium cyanates in water similar values of k_0 , E_A and A .

It is interesting to find that the order of the rates of the reactions of the three cyanates in alcohol at 10° is given by the values of a required to satisfy the Brönsted-Gronwall, La Mer and Sandved theory (p. 211), namely: ammonium cyanate, 5.0 Å (extrapolated from the 32 – 16° data); methylammonium cyanate, 7.6 Å; ethylammonium cyanate, 7.1 Å. The significance of this is not apparent.

We are indebted to Messrs Imperial Chemical Industries, Ltd., and to the Trustees of the Moray Fund for grants for the purchase of apparatus.

SUMMARY

The rates of transformation of ammonium cyanate to urea at 16 , 24 and 32°C , and of ethylammonium cyanate to ethylurea at 0 , 10 and 20°C have been found over the concentration range 0.0002 – 0.04 M in ethyl alcohol containing 1.9% of water, and compared with those for methylammonium cyanate at 0 and 10°C in the same solvent.

On the supposition that ammonium (or alkylammonium) and cyanate ions are the reactants, the applicability of the Brönsted-Gronwall, La Mer and Sandved theory to the results has been tested by assuming its validity, and calculating the values of a , the mean distance of closest approach of the ions, required to give agreement. For all three cyanates a has been found to have a high positive temperature coefficient. A simplified form of Christiansen's equation for ionic reactions is not generally satisfied at infinite dilution, whether r the distance between colliding and reacting ions is taken to be constant or set equal to the temperature-variable a values.

The Arrhenius equation ($k = Ae^{-E_A/RT}$) is not fulfilled, except at the highest concentration. E_A and A vary in a parallel manner from one cyanate to another, and with change of concentration, temperature and solvent. An

attempt has been made to explain the relative rates of transformation of the cyanates in water and ethyl alcohol at infinite dilution by supposing that the solvation of the reactants plays an important part.

REFERENCES

- Brönsted 1922 *Z. phys. Chem.* **102**, 169.
Christiansen 1924 *Z. phys. Chem.* **113**, 35.
Davies 1933 "The Conductivity of Solutions", p. 208. London: Chapman and Hall.
Gronwall, La Mer and Sandved 1928 *Phys. Z.* **29**, 358.
Harned 1938 *J. Amer. Chem. Soc.* **60**, 336.
v. Kiss and Bossányi 1935 *Z. anorg. Chem.* **224**, 33.
La Mer 1933 *J. Chem. Phys.* **1**, 289.
La Mer and Kamner 1935 *J. Amer. Chem. Soc.* **57**, 2662.
La Mer and Miller 1935 *J. Amer. Chem. Soc.* **57**, 2674.
Lowry 1934 *Trans. Faraday Soc.* **30**, 375.
Miller 1934 *Proc. Roy. Soc. A*, **145**, 288.
— 1935 *Proc. Roy. Soc. A*, **151**, 188.
Moelwyn-Hughes 1936 *Proc. Roy. Soc. A*, **155**, 308.
Ogg and Rice 1937 *J. Chem. Phys.* **5**, 140.
Percival and La Mer 1936 *J. Amer. Chem. Soc.* **58**, 2413.
Scatchard 1930 *J. Amer. Chem. Soc.* **52**, 52.
Svirbely and Warner 1935 *J. Amer. Chem. Soc.* **57**, 1883.
Svirbely and Schramm 1938 *J. Amer. Chem. Soc.* **60**, 330.
Verhoek 1934 *J. Amer. Chem. Soc.* **56**, 571.
Walker and Hambly 1895 *Trans. Chem. Soc.* **67**, 746.
Walker and Appleyard 1896 *Trans. Chem. Soc.* **69**, 193.
Warner and Stitt 1933 *J. Amer. Chem. Soc.* **55**, 4807.
-

The use of a vertical pipe as an overflow for a large tank

BY A. M. BINNIE

(Communicated by R. V. Southwell, F.R.S.—Received 30 June 1938)

[Plates 3–6]

1. INTRODUCTION

This paper describes small-scale experiments made on the flow of water from a large tank through a vertical pipe projecting upwards through the bottom for a considerable distance. Under these conditions the discharge is greatly affected by the type of flow existing at the entrance to the pipe, and at certain heads considerable volumes of air are carried down the pipe with the water.

The literature of the subject appears at present to be scanty. The classical researches of Borda (1766), which are described in text-books on hydraulics, were carried out on uniform pipes projecting inwards for a short distance only. He distinguished two types of flow from the bottom of a tank in which the water was sensibly at rest: (*a*) free flow, in which a vena contracta was formed at entrance to the pipe and the jet was steady and of the same diameter as the vena contracta; (*b*) full flow, in which the stream, after passing the vena contracta, expanded to fill the pipe and issued as a turbulent jet. Later writers—e.g. Bilton (1908), Smith and Walker (1923)—have determined experimentally the coefficients of discharge and contraction for varying pipe diameters and heads.

Mathematical hydrodynamics affords little assistance. The only similar problem that has been solved (Lamb 1932) concerns the flow in the absence of gravity of an inviscid liquid through a two-dimensional mouthpiece.

Unfortunately the artifice of conformal representation, which is employed, cannot be extended to deal with a circular pipe. The power of entraining air possessed by water falling down a vertical pipe has been used to operate hydraulic air compressors, and several large examples of plant of this kind were described by Peele (1930).

In recent years vertical pipes of bell-mouth form at entrance have been used instead of weirs as overflows for large reservoirs. W. J. E. Binnie (1937) explained the advantages of this type of overflow, reviewed examples from different parts of the world and detailed his own work on the subject. His

observations, which were carried out on large models, were in some cases checked by the performance of prototypes. These circular weirs (as they are sometimes termed) are usually operated at comparatively low heads and a vortex may be formed at inlet. His experiments demonstrated the importance of inserting radial ribs into the bell-mouth in order to minimize the effect of the vortex in reducing the flow. He also drew attention to the great volume of air sucked down by the vortex.

The experiments to be described were designed to elucidate the different types of flow which are possible in a vertical overflow pipe. Provision was made for the insertion in turn into the apparatus of pipes of various sizes and shapes, and for the investigation of the flow under heads ranging from zero to those so great that the surface of the tank was undisturbed.

2. DESCRIPTION OF APPARATUS

In order to reproduce as accurately as possible the conditions which prevail at points far from the sides of a large tank, every effort was made to ensure that the flow of water to the pipe inlet possessed no tangential component of velocity. The apparatus, which was constructed and used at the Engineering Laboratory, Oxford, is shown in fig. 1. The tank *A*, in which the pipe under test was inserted, was circular, 4 ft. in diameter and 2 ft. 6 in. in height. In the centre of the bottom a hole 5 in. in diameter was cut, and over this the brass casting *B* was bolted. The base of the casting was closed by the plate *C*, to which the pipe *D* under test was soldered. Water was supplied from a constant-level tank in the laboratory roof through a control valve to the 2 in. openings, diametrically opposite, in the casting *B*. It then passed vertically upwards in the annular space between the casting *B* and the pipe under test, any rotational velocity about a vertical axis being eliminated by a honeycomb, not shown in fig. 1. After passing through the bottom of the tank, it was constrained to flow outwards in a radial direction by the disk *E*, 2 ft. 11 in. in diameter, which was supported on three small feet, $1\frac{3}{4}$ in. long and symmetrically arranged. The clearance between the disk and the pipe under test was sealed by a brass plate and a rubber ring, and thus the incoming water was prevented from disturbing the flow at the pipe inlet.

The arrangement described above proved most successful in ensuring that the water reached the pipe inlet radially when seen from above. Floating particles dropped on the surface in the tank were invariably observed to travel in straight lines direct to the pipe inlet except when, as mentioned later, a vortex formed near the centre of the tank.

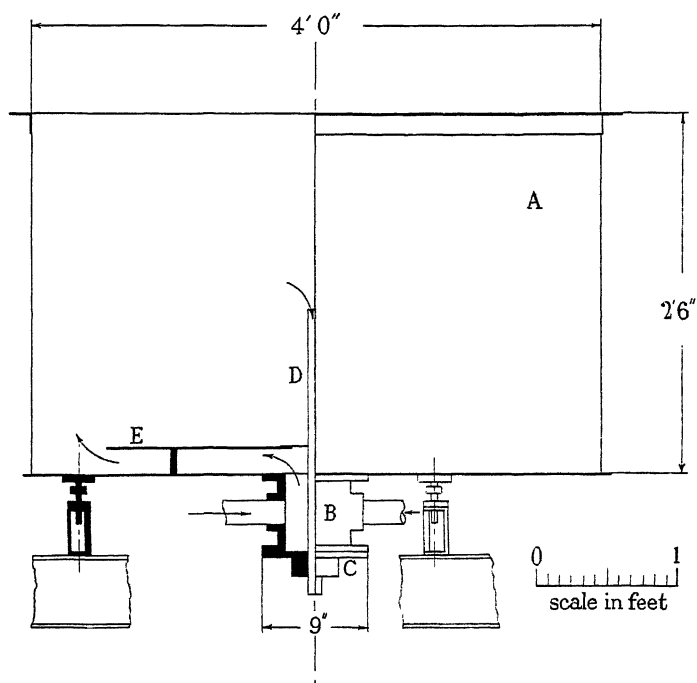


FIG. 1. Elevation of apparatus.

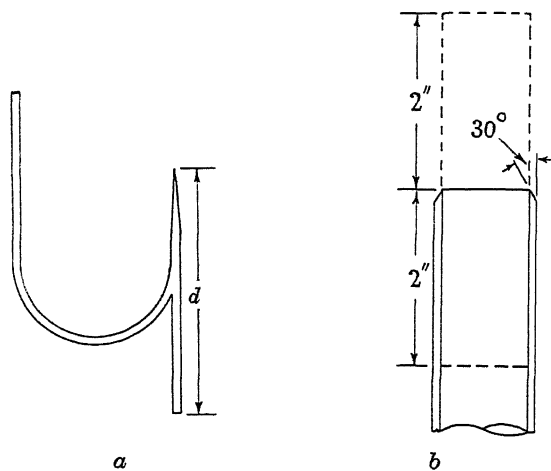


FIG. 2. *a*, Subsidiary hook gauge. *b*, Sharp ended 1 in. pipe with central plate.

After passing through the pipe under test the water fell into a calibrated measuring tank. The level in the tank *A* was found by means of a hook-gauge placed near the circumference of the tank on a cross-bar. The effects of surface tension at the pipe inlet prevented the direct reading of the zero of this gauge. A subsidiary hook-gauge with its end formed to the shape shown in fig. 2*a* was therefore mounted centrally on the cross-bar. The lower end of the hook was sharp or flat according as the end of the pipe under test was flat or sharp. With this gauge, two readings were taken, the first when the lower point was touching the top of the pipe. The lower end of the pipe was then closed by a cork and the water level in the tank raised a couple of inches. The second reading was made with the upper point in the surface, and the main hook-gauge was observed at the same time. The distance *d* between the two points of the gauge was known, and thus the desired zero was easily found. The subsidiary gauge, when not in use, was drawn up out of the water.

3. TESTS ON A 1 IN. UNIFORM PIPE 2 FT. IN LENGTH

The object of these tests was to determine the relation between the discharge *Q* and the head *H* of water above the top of the pipe. The procedure was to adjust the control valve in the supply pipe, and when conditions had become steady to measure simultaneously the discharge and the level of the surface in the tank.

The first pipe to be tested was a stainless steel tube 2 ft. long, $1\frac{1}{8}$ in. outside diameter and wall thickness 0.064 in. (16 gauge). It was used, as supplied by the makers, "in the fully softened and descaled condition". Its top end, which was flat, was $11\frac{3}{4}$ in. above the disk baffle, so that the effects of the tank bottom upon the flow may be considered negligible. Its bottom end just projected below the brass casting.

In the first series of observations the opening of the control valve was successively increased by small amounts, producing what may be termed "normal rising flow". The results are given by the curve on the right of fig. 3: they were not all obtained on the same day, and their consistency indicated that there was no difficulty in obtaining repeatable results. The test was commenced at zero head. Owing to the effects of surface tension, there was no flow until the head was about 0.2 in., when the water flowed steadily down in contact with the surface of the pipe, leaving a central hole of air. At exit it formed an annular jet of glassy appearance, which in the course of a fall of 2 or 3 in. was pulled by surface tension to the pipe axis and broken up there. The flow being increased, the central hole was reduced

in size and the annular jet lost its glassy appearance. At a head of about 0.45 in. the type of flow described above became unstable, and the central hole opened and shut at about $\frac{1}{2}$ sec. intervals, producing a "gulping" noise and causing ripples in the tank. The jet at exit was disturbed in a similar periodic manner. The head being increased to 0.5 in., the sound produced from the central hole rose rapidly in frequency to a low snarling noise; the jet remained irregular but the surface of the tank was undisturbed. At a

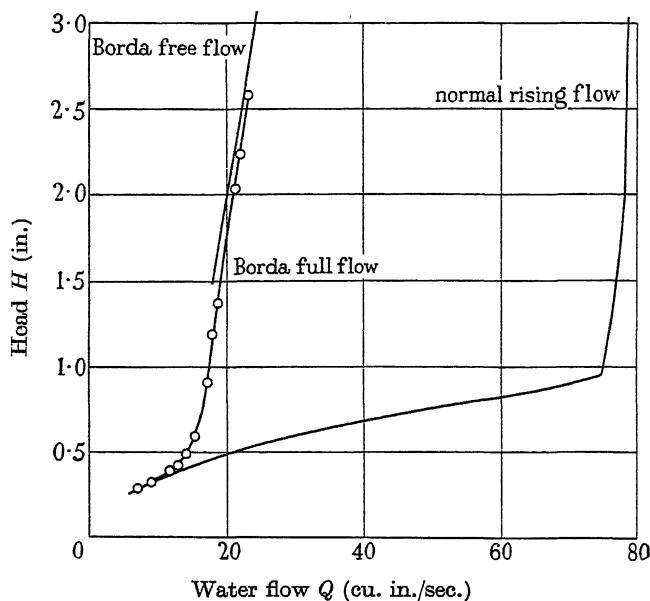


FIG. 3. Tests of the flat-ended 1 in. uniform pipe, 2 ft. in length.

head of about 0.75 in. the pitch of the noise emitted reached its highest value, and the jet became regular. A further increase of head caused the noise to be choked.

At a head of approximately 0.93 in. a very marked change in the flow took place. The central hole filled in, assuming a "fluttering" form which caused ripples in the tank. Up to this point the flow in the tank appeared to be radial, and the motion of dust particles on the surface had no rotational component. It was difficult to see what was occurring at the pipe entrance, and therefore further experiments, to be described later, were carried out to elucidate this point. A slight increase of head above this stage caused the continuous central hole to disappear and the level in the tank rose enormously. As will be seen from fig. 3, a critical head was reached at

0.95 in., at higher heads the increase of water flow being very small. At first, when the head did not greatly exceed this value, a single circular vortex with an air core was intermittently formed, accompanied by considerable noise. But it usually persisted only for a short time and failed to cause any rotation of the surface of the tank outside a radius of a couple of inches. Above a head of 3 in. these vortices were rare: they sometimes attempted to form but did not develop into more than a small dimple on the surface.

It should be noted that except in the brief stages of gulping and fluttering the surface of the tank outside a radius of a couple of inches remained entirely steady. The ripples set up at the former stage were small, though occasionally gulping was sufficiently violent to project drops of water to a height of 2 or 3 in. Fluttering at the critical head was rather more serious; circular ripples were formed and the level of the surface was not absolutely steady. Thus at this point the head could not be measured with quite the same precision which was easily obtainable at all other heads. On the steeply rising portion of the curve the tests were extremely tedious. The slightest opening of the control valve caused such a great increase in the head that conditions did not become steady until the lapse of an hour or more.

A series of readings was then obtained for falling values of the head. The tests were begun at the highest measurable head and the control valve was progressively closed. It might have been thought that the curve obtained would show some tendency to cut the sharp corner of the normal rising flow curve in the region of the critical head. No trace of this, however, could be detected; the falling curve was throughout the whole range identical with the rising.

On two occasions, when normal flow at a head of about 2.5 in. was under investigation, the flow changed suddenly of its own accord to Borda free flow. The issuing stream assumed the appearance of a thin fixed glass rod projecting from the centre of the pipe but not touching it, the noise of impact of the jet in the measuring tank ceased, and the level in the tank *A* immediately rose. It was therefore necessary to investigate this type of flow, and a method of producing it at will was devised. A cork (to which a rod was attached as a handle) was inserted at the top end of the pipe and the head in the tank was raised to about a couple of inches. The cork was then carefully withdrawn by hand in a truly vertical direction. About half the attempts resulted in the immediate production of Borda free flow.

This type of flow was found to be possible under heads ranging from the greatest value which could be measured down to a lower limit of about

1.5 in. This lower limit varied slightly in different tests. The observations of head and discharge are given on the left of fig. 3, whence it will be seen that, for any head, the discharge compared with normal flow is very greatly reduced. Throughout this series of tests the surface in the tank remained flat and undisturbed except on one occasion when the head was slightly greater than the lower limit. A noisy vortex was then formed with an air core which extended down through the glassy jet and was visible in the jet underneath the tank; this caused the jet to break up into drops about 2 ft. below the bottom of the tank. Borda free flow was found to be surprisingly stable. The water in the tank could be agitated with sufficient violence to break up the jet so that a splash was momentarily heard in the measuring tank, yet the jet would reform itself and soon settle down into a state of complete quiescence.

As explained above, there was a lower limit of head below which Borda free flow was impossible. If the head was reduced below this limit the type of flow known as Borda full flow occurred of its own accord. Here the stream after contracting at entrance to the pipe expanded to fill the pipe and issued as an extremely turbulent jet. The surface in the tank remained undisturbed and flat. The observations on this type of flow are also shown on fig. 3. They were continued down to very low heads, when the central hole appeared and conditions were similar to those observed in the initial stages of normal rising flow. They were not, however, quite identical because a small ring of air was seen to be entrapped round the inner edge of the top of the pipe. This was not present with normal flow.

Observations of Borda full flow were also made with increasing values of the head. Full flow was first obtained by reducing free flow below its lower limit and then the control valve was progressively opened. It was found that the head could be increased to about 2.5 in. before the flow changed of its own accord to free flow. Thus there is an overlap of about 1 in. of head between the curves of free and of full flow. Between the limits 2.5 and 1.5 in. of head both types of flow were possible; which type occurred depended on the previous history of the flow.

It is not difficult to understand why Borda flow at any head caused such a great reduction in the discharge as compared with normal flow. Under free flow the discharge was independent of the length of the pipe and was small because the diameter of the jet was small. Under full flow the available potential energy of the stream was dissipated in a jet which was observed to be exceptionally turbulent, so unsteady that drops of water were flung out laterally and descended into the measuring tank clear of the main jet.

The top end of the pipe was then sharpened (fig. 2*b*) and the experiments described above were repeated leading to results (fig. 4) which are the same in general form as those obtained for the flat-ended pipe. The accompanying phenomena of noise, vortices, etc., were the same as before. The normal rising flow curve is almost identical with that determined for the flat-ended pipe, except that the critical head was reached at about 0.92 in. instead of 0.95 in.; hence the steep portion of this curve lies to the left of that previously obtained. Again, there was no appreciable difference between rising

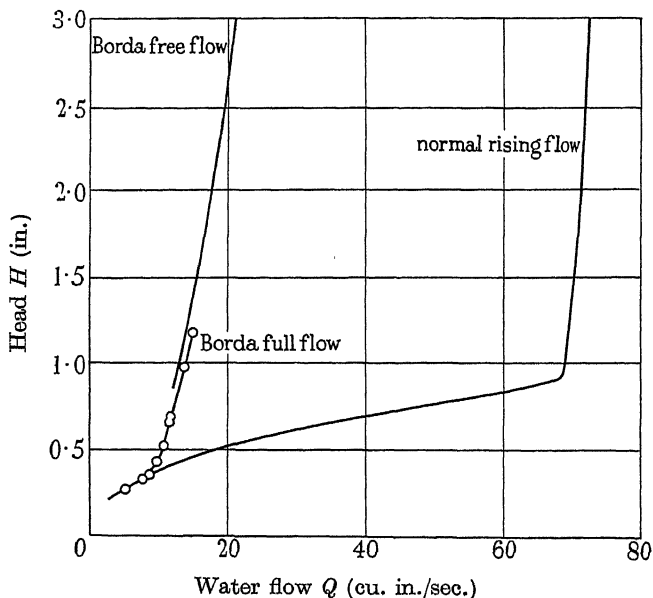


FIG. 4. Tests of the sharp-ended 1 in. pipe, 2 ft. in length.

and falling normal flow. The Borda flow curves are also moved to the left. Borda free flow was possible down to a head as low as about 0.8 in., and full flow up to a head of about 1.2 in. Thus free flow was facilitated by sharpening the end of the pipe, the zone of overlap being reduced and brought much lower.

An explanation of the steepness of the normal curve above the critical head may be attempted on the assumption that the pipe ran full throughout its length. It follows that the gross head producing flow was $(24 + H)$ inches and that in the absence of losses the flow Q should vary as $(24 + H)^{\frac{3}{2}}$. Now the velocity through the pipe and therefore the losses in and at entrance to the pipe did not greatly alter with the head, hence for the present purpose the coefficient of discharge may be regarded as constant and thus may be

neglected in this comparison. This view may be tested by analysing eight tests conducted at such high heads that disturbances were rare. The necessary figures are set out in the table below, the last line but one in which shows the flow Q in tests (ii)–(viii) calculated on the supposition that it could be predicted from test (i) in the manner described above. The agreement between the observed and calculated values of Q is seen to be as close as could be expected.

SHARP-ENDED 1 IN. UNIFORM PIPE. INCH-SECOND UNITS

Test	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
H (observed)	2.17	2.66	3.10	3.31 ₅	3.99	4.26	5.07 ₅	5.83
$(24 + H)^{\frac{1}{2}}$	5.116	5.163	5.206	5.226	5.291	5.316	5.392	5.462
Q (calculated)	—	72.1	72.8	73.1	73.9	74.3	75.4	76.4
Q (observed)	71.5	71.9 ₅	72.4	72.7	73.8 ₅	74.0	75.4	76.4

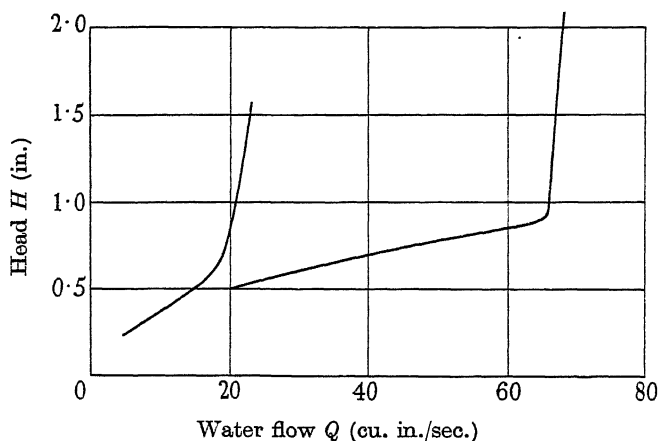


FIG. 5. Tests of the sharp-ended 1 in. uniform pipe, 2 ft. in length, with central plate.

As mentioned above, it was uncertain whether the normal flow at entrance to the pipe under heads below the critical was rotational or not. The motion of the surface surrounding the pipe was seen to be radial but it was impossible to observe the flow inside the pipe itself. To throw light on this question, a sheet of brass, 0.0345 in. thick, 4 in. long and 1 in. wide, was jammed vertically and symmetrically in the sharp-ended pipe (as shown in fig. 2*b*), so that it protruded 2 in. into the pipe and projected 2 in. above it. The cross-section of the pipe was thus divided into two equal portions. Observations of head and discharge were then conducted in the usual manner. The results, which are displayed in fig. 5, were unexpected.

When the test was carried out in the customary way with small increases of the head, at first the quiet central hole was formed as before, although now divided by the plate. A further increase of head caused the hole to fill in, the surface of the tank was flat and undisturbed, and the flow was noiseless. This type of flow persisted up to the maximum head of $1\frac{1}{2}$ in. which was measured, the observations lying on the steeply rising branch of the curve on the left of fig. 5. The flow appeared to be of the Borda full type, modified somewhat because the curve lies further to the right of that obtained without the plate than can be attributed to the small reduction in cross-section due to the plate. This view was supported by the exceptionally turbulent nature of the jet, which closely resembled that produced by Borda full flow without the plate.

If, however, the head was raised suddenly from zero, the familiar noisy normal flow with a central hole set in, giving the flat branch of the curve on the right of fig. 5 which was identical with that obtained without the plate. The central hole was seen to be divided symmetrically by the plate. A critical head was reached at much the same value as before.

It is therefore clear that, in the unobstructed pipe, normal flow below the critical head was not rotational. The noise was produced by a high-frequency form of gulping, the central hole opening and shutting with great rapidity. It was this action which caused considerable volumes of air to be carried down the pipe.

An attempt was made to determine the air flow by placing a bell jar over the top of the pipe and connecting it to a measuring device. But no measurement of air flow could be carried out without causing a small pressure variation in the bell, and this greatly disturbed the water flow. The air measurements had consequently to be carried out at the lower end of the pipe. The 2 ft. pipe projected below the brass casting for too short a distance for any air-measuring device to be connected to it. It was therefore later replaced by a longer pipe which gave ample space for the necessary apparatus. The experiments with this are described in the next section.

In order that the air flow down the pipe might be studied, the steel pipe was replaced by a flat-ended glass tube as nearly as possible of the same dimensions. A camera, directed vertically downwards, was fixed on the cross-bar near the edge of the tank *A*, and beneath it a large flat mirror was held submerged in the tank at an angle of 45° to the horizontal. In this way, views of the glass tube were recorded similar to those which would be obtained by an observer stationed with his eye under water and looking horizontally. The photographs, which are reproduced in fig. 6, Plate 3, indicate clearly the different types of flow. At a very low head a vena contracta was

formed, and the water adhered to the pipe, leaving an undisturbed central hole of air. This is shown in (a), which can be interpreted when it is noted that the pipe filled and surrounded by water appears white except at the edge. When filled with air but surrounded by water it appears black except for a central band of white. The inverted image at the top of the photograph is due to reflexion from the under side of the water surface. The distance between the top of the pipe and its image gives an indication of the prevailing head, and the water surface lies between them. This photograph was taken with an exposure of 3 sec. Various stages in the process of gulping at constant head are shown in *b* (i), (ii) and (iii), which (like the remainder of the photographs) were given an exposure of 1/500 sec. These demonstrate that the air was taken down by blocks of water almost completely isolated from one another. Photographs (c), (d) and (e) show the flow at successively greater heads; in (c) flow attended by low frequency noise had just commenced, (d) was an intermediate stage, and in (e) the noise had begun to die away. These photographs indicate that what occurred was a high-frequency form of gulping, taking place with increasing rapidity. Fluttering and the accompanying ripples are shown in (f), a transient vortex in (g), and the base of a high-head transient vortex in (h).

4. TESTS ON A 1 IN. UNIFORM PIPE, 5 FT. IN LENGTH

As mentioned in the previous section, the 2 ft. pipe was so short that air-measuring apparatus could not be attached to its lower end. It was therefore replaced by a pipe of the same material and diameter but 5 ft. in length. This was fixed in the tank so that its top, which was sharpened as shown in fig. 2*b*, occupied the same position as before. The air-measuring device (fig. 7) was arranged at the lower end of the pipe. The water after leaving the pipe fell into the open tank *P*, one edge of which was cut away to form a long weir. Thus the level in this tank remained almost constant whatever the flow. After passing over the weir, the water dropped into a calibrated measuring tank. The air which passed down the pipe with the water was trapped in the bell *Q*, 19½ in. high and measuring 12 by 12 in. in plan. The weight of the bell was carried by four adjustable feet resting on the bottom of the tank *P*, and a gland was provided where the pipe under test passed through the top of the bell. The air pressure in the bell was measured by the water U-gauge *R*. The air was led from the top of the bell through the large vessel *S*, which served to steady the flow, to the calibrated gas meter *T*. To overcome the frictional resistance of the piping and the

meter, the ejector fan V was connected to the meter outlet; and by adjusting this fan, the air pressure in the bell could be kept steady at any value. It was found necessary to hang the plate X below the end of the pipe in order to break the fall of the water.

It was desired (as in the previous tests) to measure simultaneously the head and the water flow and at the same time to determine the air flow

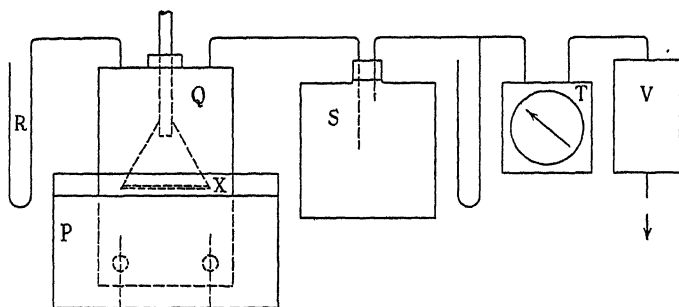


FIG. 7. Air-measuring apparatus.

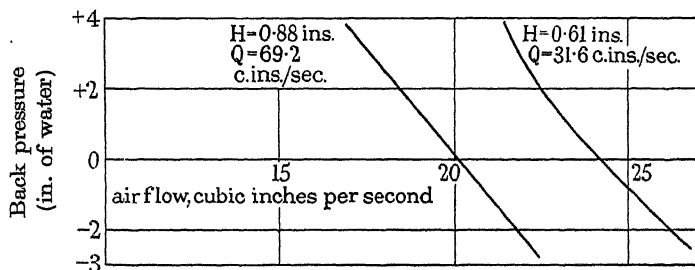


FIG. 8. Variation of air flow with back pressure.

with the pressure in the bell exactly atmospheric. But before this could be done with confidence, it was essential to show that the air flow was not altered by undetectable variations in the back pressure in the bell. Two tests were therefore carried out with constant water flow, and in each the head and the air flow were measured when the back pressure was varied between wide limits. The head was found to remain constant, except for a just perceptible rise of 0.1 mm. when the back pressure exceeded 2 in. of water. The variation of air flow with back pressure for the two tests is shown in fig. 8, from which it will be seen that the air flow was comparatively insensitive to the back pressure. Without difficulty the back pressure could be maintained atmospheric within $1/40$ in. of water, and it was therefore

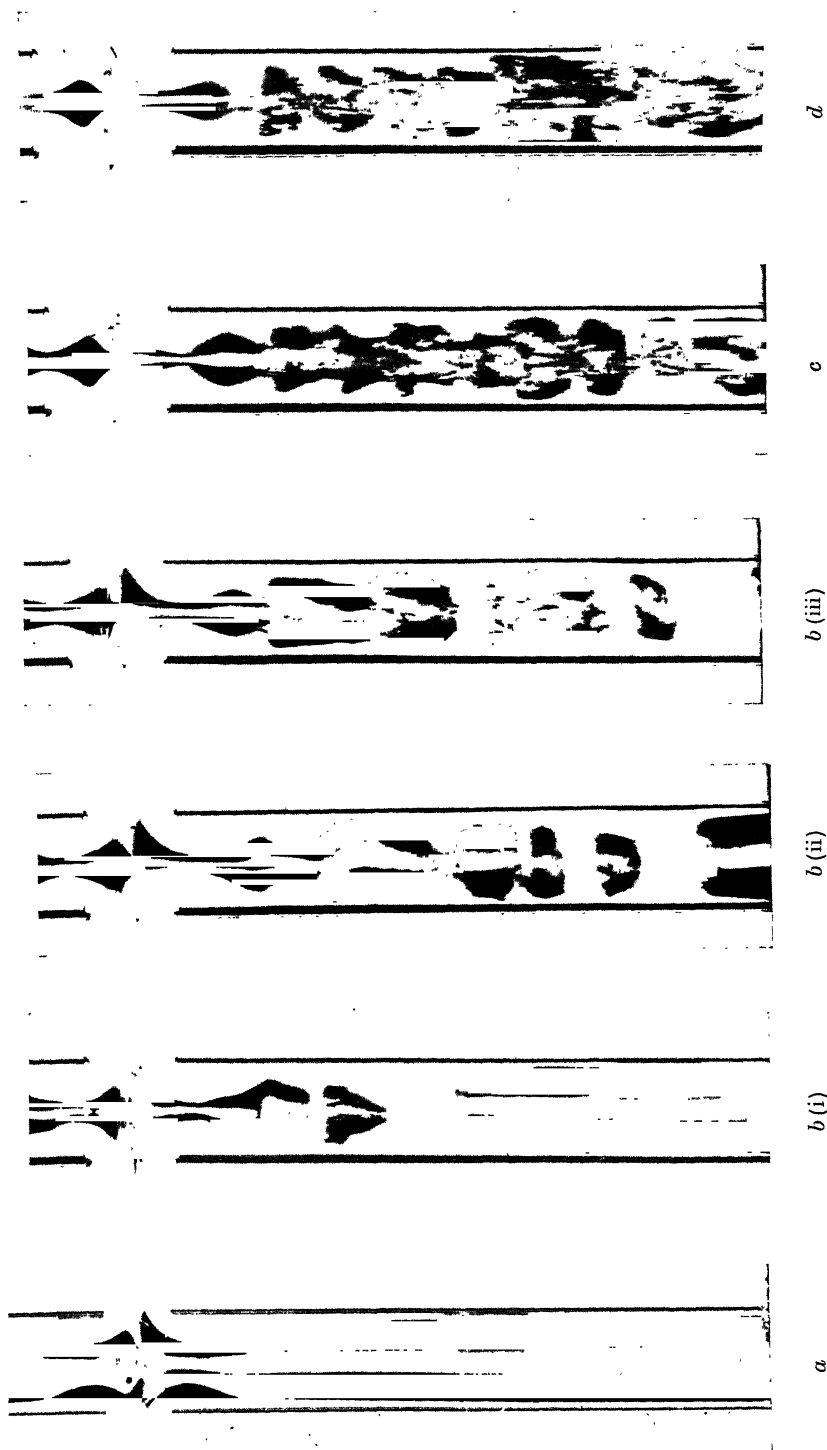


FIG. 6. Photographs of the flow through the flat-ended pipe.

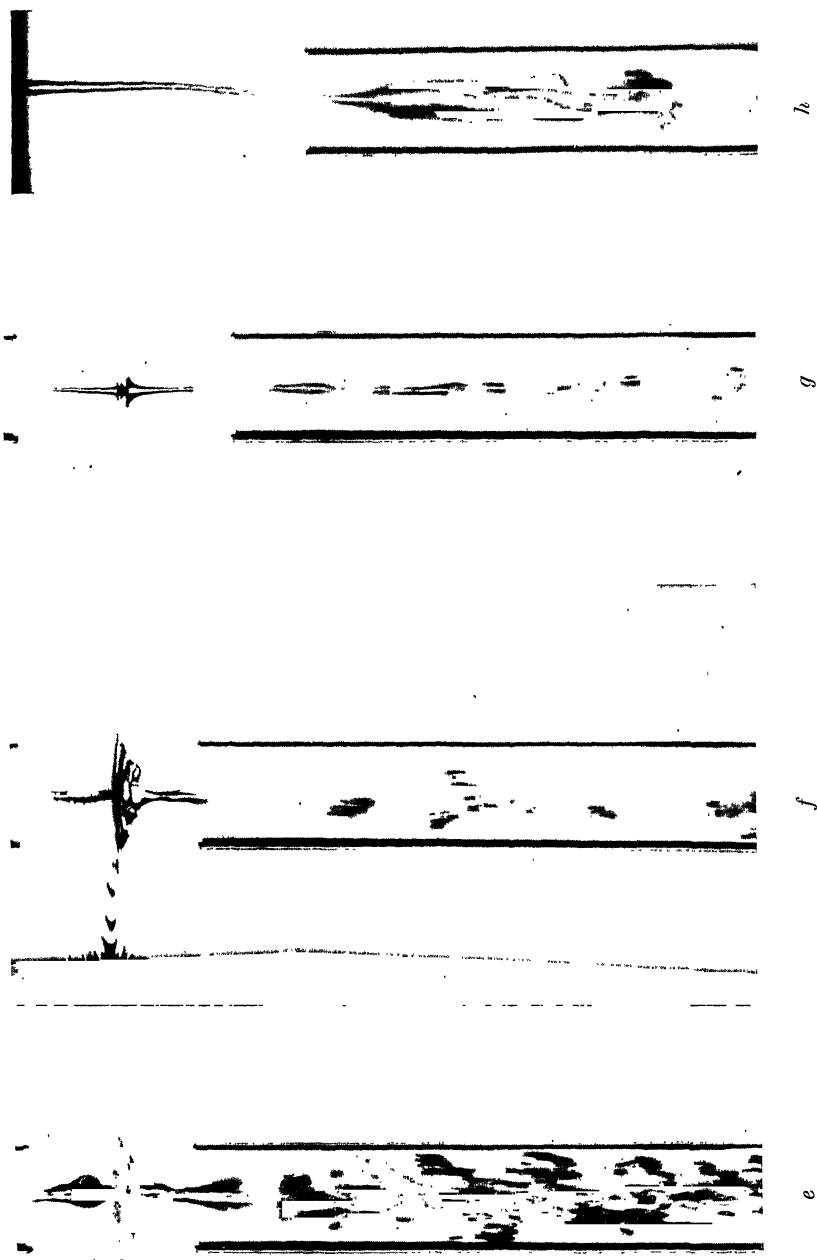


FIG. 6 (continued)

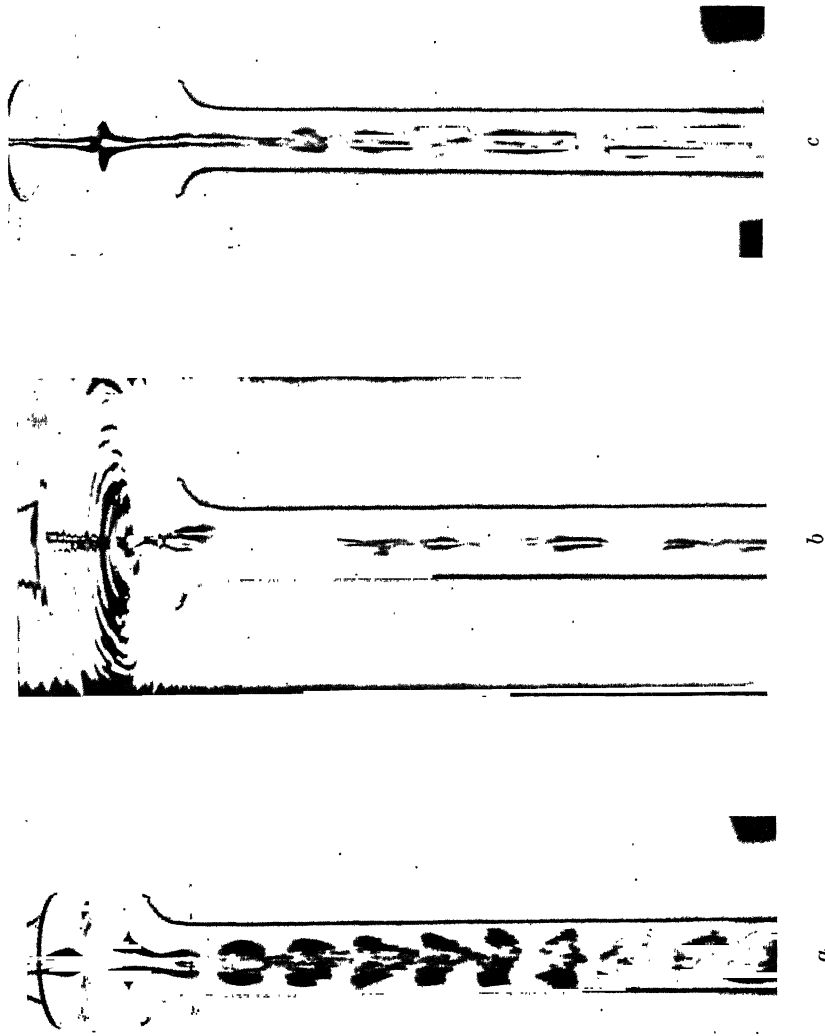


FIG. 12. Photographs of the flow through the pipe with a short rounded entrance.

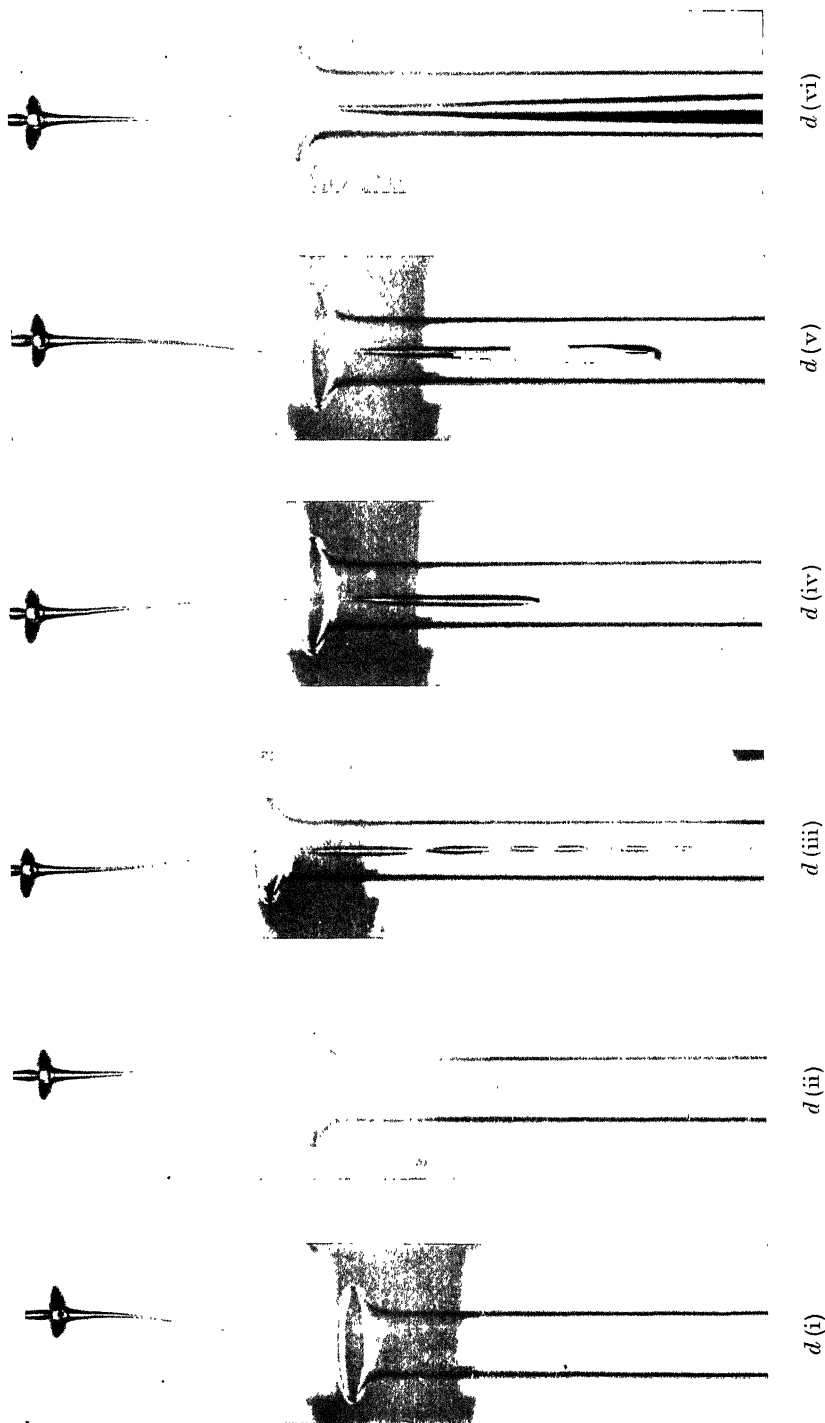


FIG. 12 (continued)

concluded that the proposed method of measuring the air flow was satisfactory.

Tests on normal rising flow were then performed, yielding the results shown in fig. 9, where, on a common base of water flow Q , the values of the head H , the air flow and the volumetric ratio R of air flow to water flow are plotted. The effects on the water flow of increasing the pipe length from 2 to 5 ft. may be seen by comparing these results with fig. 4. At low heads there was no change in the readings, but with the longer pipe the flat portion of the curve persisted further, the critical head being increased to

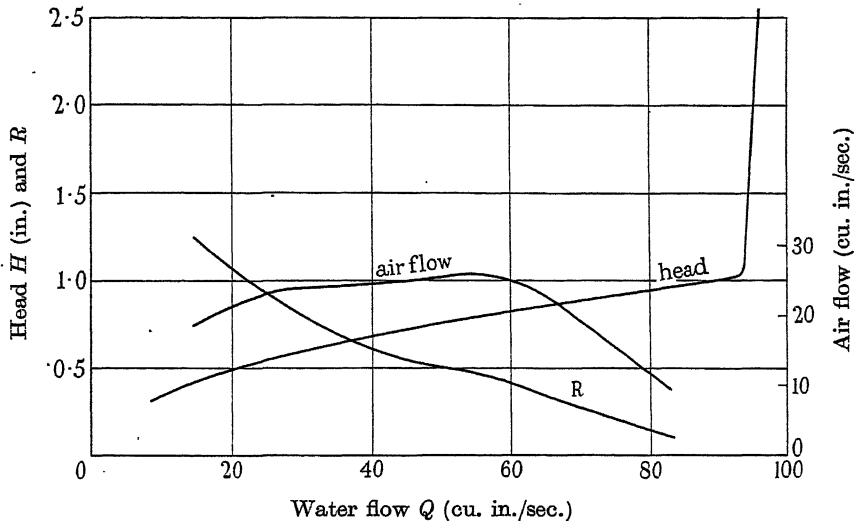


FIG. 9. Tests of the sharp-ended 1 in. uniform pipe, 5 ft. in length.
 R = air flow/water flow.

about 1.02 in. It was found impossible to obtain a satisfactory reading when the head was slightly greater than the critical. The transient vortices which were formed were sufficiently powerful to obstruct the water flow; and although the water supply to the tank A was kept constant, the head H was seen to surge by as much as $\frac{1}{4}$ in. Air flow suddenly began when gulping commenced at a head of 0.42 in., and instantly attained the considerable magnitude of 18.3 cu. in./sec. It increased slowly to a maximum of 26.0 cu. in./sec. at a head of 0.785 in. and a water flow of 54.4 cu. in./sec., and then fell away to zero just before the critical head was reached. The vortices formed above the critical head were so transient that the air flow induced by them could not be measured.

5. TESTS ON A 1 IN. UNIFORM PIPE FITTED WITH A TRUMPET-SHAPED ENTRANCE 5 IN. IN DIAMETER

To discover what improvement could be effected by fitting a trumpet-shaped entrance to the pipe, the top of the 1 in. uniform pipe (mentioned in the previous section) was cut off and replaced by the conical trumpet shown in fig. 10. This was a machined brass casting, geometrically similar to that shown by full lines in fig. 9, Plate 2, by W. J. E. Binnie (1937). Its top edge was placed at the same height above the bottom of the tank as that of the pipes previously tested, and the overall length of the pipe and trumpet remained 5 ft. Tests, similar to those described in the previous section, were then carried out in this arrangement, the results being plotted in fig. 11.

At low heads, up to 0.35 in., the water ran quietly down the surface of the trumpet and pipe, leaving a central hole of air. The only noise was that due to splashing at the lower end of the pipe. At this stage the air flow was considerable but was extremely sensitive to the back pressure. The water U-gauge was therefore replaced by a differential U-gauge, containing benzyl alcohol and a solution in water of calcium chloride, which gave a magnification of 7.95. The air flow at zero back pressure for various heads was then deduced from curves of the same form as those shown in fig. 8. The air flow was thus found to range from 24.5 cu. in./sec. when the head was 0.18 in. to 17.4 cu. in./sec. at a head of 0.30 in., but these results have not been added to fig. 11 because they were so sensitive to minute changes in the back pressure. At a head of 0.36 in. and a water flow of 49.0 cu. in./sec., low-frequency gulping occurred, the water apparently running down the surface of the trumpet with sufficient velocity to leap clear of the pipe into the central hole. The air flow rose to 41.4 cu. in./sec., which was the maximum observed, both absolutely and relatively to the water flow. At this and higher heads the air flow was found to be insensitive to the back pressure. A slight increase of head to 0.38 in. caused the gulping to increase in frequency, a loud noise was produced, doubtless accentuated by vibrations of the trumpet itself, and the air flow fell to 35.1 cu. in./sec. Further rises of head caused the noise first to increase in violence and frequency and then to die away as the critical head was approached. This was reached at about 0.60 in. and was as sharply marked as in the tests on uniform pipes. At a head of 0.59 in., there was still a small air flow but signs of incipient gorging were apparent in slight splashing at the bottom of the trumpet. When the head was increased by 0.1 mm., the smallest change which could be measured, the water level in the trumpet rose to within a couple of inches of the top

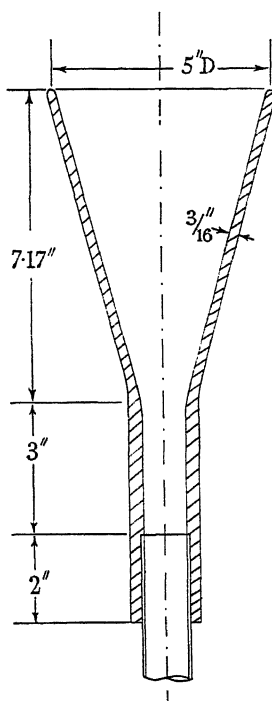


FIG. 10. The 5 in. trumpet-shaped entrance to the 1 in. uniform pipe.

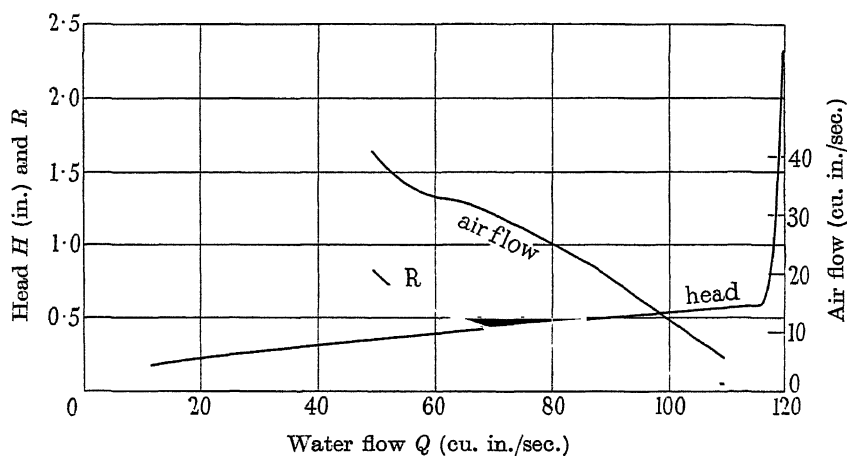


FIG. 11. Tests of the 1 in. uniform pipe with 5 in. trumpet-shaped entrance.
 R = air flow/water flow.

edge. Its surface was agitated but there were no signs of the stable vortices and vertical surges described by W. J. E. Binnie (1937, photographs nos. 5, 6 and 7). An increase of 1 cu. in./sec. in the water flow caused the head to rise to 0.72 in. when the trumpet was completely submerged; a vortex feebly attempted to form but was defeated by fluttering. At greater heads the water surface was undisturbed except for small dimples which never quite developed into vortices. At heads above the critical, small variations in the head were observed when the water flow was kept constant, but they were less marked than those which were noticed when the 1 in. uniform pipe was under test. No doubt they were caused by the dimples mentioned above, which (although small) could cause considerable disturbance to the head at this stage when the variation of head with water flow was so great.

A comparison of fig. 11 with fig. 9 shows that a considerable but not immense improvement in the performance of the overflow was effected by the trumpet-shaped entrance. The water flow at the critical head was increased from 93 to 116 cu. in./sec. a gain of 26 %, but, perhaps more significantly, the critical head was reduced from 1.02 to 0.60 in. The air flow was much increased, attaining its maximum value at almost the same water flow as before. This point, however, does not appear to be of great importance because this maximum air flow occurred when the water flow was less than half its value at the critical head.

At various heads attempts were made by stirring the water in the tank to initiate a stable vortex, but all were unsuccessful. It appears possible, therefore, that this can only be produced when, as in reservoir practice, an inclined tail piece is attached to the vertical portion of the pipe. Alternatively, the absence of a vortex may be due to some unexplained property of this shape of entrance, which was found by W. J. E. Binnie (1937, §D2) to be particularly satisfactory. The only point of interest which was observed was that, when the trumpet was just awash, waves which broke over it caused a considerable volume of air to be trapped round the inner side of its top edge.

6. PHOTOGRAPHS OF THE FLOW THROUGH A GLASS PIPE FITTED WITH A TRUMPET-SHAPED ENTRANCE

It was found impracticable to make in glass a replica of the pipe with a trumpet-shaped entrance described in the previous section, but some photographs were taken (at 1/500 sec. exposure) of the glass pipe previously examined when its top end had been drawn out to form a short rounded entrance. The photographs are reproduced in fig. 12, Plate 5, which also

indicates the shape of the entrance. This type of entrance produced at high heads an exceptionally violent form of transient vortex, affording an opportunity of examining in detail this kind of disturbance. Noisy flow at a low head is seen in (a), which shows clearly the regular sequence of the air bubbles passing down the pipe. Fluttering (with its accompanying ripples) near the critical head is shown in (b), and a thick low-head vortex in (c). The remaining photographs (d), which were taken at the same high head, disclose successive stages in the formation of a transient vortex. The first sign of disturbance was a dimple (*d. i*), which ejected small bubbles of air into the pipe. This developed into a thin column (*d. ii*) of slightly irregular diameter. The swellings then increased in size (*d. iii*) and, if growth persisted, amalgamated into one large cavity (*d. iv*). This in turn became larger (*d. v*) and finally (*d. vi*) reached down the tube, emitting a harsh and violent noise and drawing down a considerable volume of air. The process of dying away was usually much more rapid; the vortex decreased in diameter and quickly vanished.

7. CONCLUSIONS

The results of the experiments may be summarized as follows:

(i) Through the four arrangements which were tested, normal water flow increasing from zero caused at first only a small rise of head. This stage was accompanied by the noisy entrainment of air bubbles. Above a certain critical head the increase of water flow with head almost ceased. The values of these critical heads and the corresponding flows of water are given in the table below:

Overflow	Critical head in.	Water flow cu. in./sec.
(a) Flat-ended 1 in. pipe, 2 ft. in length	0.95	75
(b) Sharp-ended 1 in. pipe, 2 ft. in length	0.92	68
(c) Sharp-ended 1 in. pipe, 5 ft. in length	1.02	93
(d) 1 in. pipe with 5 in. trumpet, 5 ft. in overall length	0.60	116

The relation between the head and the water flow was found to be the same whether the head was rising or falling. At heads slightly greater than the critical the flow was sometimes disturbed by transient vortices.

(ii) At heads below the critical, the maximum air flow with overflow (c) was 26.0 cu. in./sec. and with (d) 41.4 cu. in./sec.

(iii) With overflows (a) and (b), Borda free and full flows were found greatly to reduce the water flow.

How far these results may be used to predict the performance of larger pipes cannot be decided with certainty. In these experiments the surface exposed in the pipes to high-velocity water was comparatively small and viscosity is likely to have had no great effect. But the principle of similarity demands also that the consequences of surface tension should be negligible, and here (apart from its influence at extremely low heads) surface tension was perhaps important in controlling the size of the air bubbles entrained by the water. Since the flow was found to be comparatively insensitive to the back pressure, it seems probable that the addition of a short inclined tail piece of large diameter would not have caused much change in the flow. But, if the tail piece offered serious resistance to the passage of air and water, the flow would be radically altered. These points could not be studied experimentally at Oxford, since the necessary facilities of space and water were not available. It is to be hoped that similar experiments, but on a much larger scale, may be made elsewhere to elucidate them.

The author wishes to express his thanks to his colleagues Mr V. Belfield, M.A., and Mr A. N. Black, M.A., for taking the photographs; to Mr D. G. Christopherson, B.A., for assistance in measuring the air flow; and to the Gas, Light and Coke Company for the loan of a calibrated gas meter.

SUMMARY

The performance of vertical pipes arranged as overflows in a tank was studied experimentally on a small scale with special apparatus, which ensured that the water reached the pipe inlet with no tangential component of velocity.

Under normal conditions, the change of head with discharge was small at low heads. At this stage the flow was not rotational, but a considerable volume of air was drawn down the pipe in the form of bubbles. Above a sharply marked critical head, the pipe ran full and a large rise of head caused only a slight increase in the discharge. The effects on the critical head of lengthening the pipe and of sharpening its inlet end were comparatively small, but the insertion of a trumpet-shaped mouthpiece greatly improved the performance.

The types of flow described by Borda as free and full flow were possible with overflow pipes of uniform diameter, and they resulted in a serious reduction in the discharge.

REFERENCES

- Bilton 1908 *Victorian Inst. Engrs*, 9, 27.
Binnie, W. J. E. 1937 *Trans. Instn Wat. Engrs*, 42, 103.
Lamb 1932 "Hydrodynamics", p. 96. Camb. Univ. Press.
Peele 1930 "Compressed Air Plant", p. 248. New York: Wiley.
Smith and Walker 1923 *Proc. Inst. Mech. Engrs*, p. 23.
-

Specific heat-temperature curves of some age-hardening alloys

BY N. SWINDELLS, B.A., AND C. SYKES, D.Sc.

(Communicated by W. L. Bragg, F.R.S.—Received 5 July 1938)

1. INTRODUCTION

In a recent paper Gayler (1937) has dealt with the various theories of age-hardening and has put forward the view that ageing takes place in two stages: in the first stage the solute atoms diffuse to planes about which precipitation proper will ultimately take place, and this gives rise to an increase in resistance to deformation and an increase in electrical resistivity without change in lattice parameter. The second stage follows the first and takes place nearly simultaneously. Some of the diffusing atoms will form molecules with neighbouring atoms of the solvent or themselves, and these molecules gradually form groups which will tend to produce a gradual decrease in resistivity and a diminution in the rate of hardening. When the molecular groups have grown to such an extent that the parent solid solution can no longer withstand the stresses set up, release of these stresses is caused by rejection of the groups, i.e. precipitation proper takes place. Once precipitation has set in then, according to Gayler's view, softening should begin.

This theory tends to combine the "Knot" theory (Gayler and Preston 1932), which states that age-hardening occurs due to the formation of clusters or groups of atoms inside the parent lattice (the first stage, according to Gayler), and the precipitation theory (Merica and others 1919) which attributes hardening to the presence of large numbers of precipitated

particles,* probably ultramicroscopic in size, dispersed throughout the material. Naturally the formation of clusters or knots will occur at a lower temperature than precipitation, so that at low temperatures hardening will occur by the mechanism of the first stage. At high temperatures hardening may take place due to precipitation, the first stage being masked entirely. The question as to which of the two processes is likely to be more effective in any particular alloy when heat-treated to give the maximum hardness obtainable in the hardening range is not specifically dealt with by Gayler, although it is stated that the experimental evidence available suggests that the second stage is, in general, more important. The work of Cohen (1936) on the silver-copper alloy containing 7.5 % copper supports this contention, since his results show that the hardening due to knot formation even at low temperatures (100–150° C) is very small indeed.

Experimental verification of any theory of age-hardening is difficult; any knot or molecular group is not readily detected by the X-ray methods at present available, and it is generally agreed that the size of the precipitated crystallites corresponding to maximum hardness is too small to be visible by normal microscopic examination. Consequently the location of the solute atoms in an age-hardened material has had to be settled by inference, a state of affairs which is not entirely satisfactory.

When precipitation takes place from a supersaturated solid solution heat is evolved, and it appeared that if sufficiently sensitive measurements of the heat of precipitation could be made, they could be used as an additional criterion of the extent to which precipitation had occurred in any material after a given heat treatment.

One of the authors (N. S.), working at Cambridge under the direction of D. Stockdale, attempted to follow the changes occurring during age-hardening by recording thermal changes in copper-aluminium and beryllium-copper alloys.† The experimental results were definitely promising and it was thought worth while to apply to the problem the technique developed for the investigation of the thermal effects associated with order-disorder transformations. As one of us (C. S.) had the necessary apparatus the present collaboration was arranged. It was realized that the investigation should include a study of a number of different age-hardening alloys in view of the probability that the changes occurring vary considerably from case to case.

* Particles having a lattice parameter typical of the second phase and not the parent solid solution.

† Using the apparatus of Quinney and Taylor (1937) kindly lent by Professor G. I. Taylor, F.R.S.

The present paper records thermal and hardness measurements on five typical age-hardening alloys. No elaborate examination of any particular alloy has been attempted; the aim has been to determine the state of precipitation in each alloy when hardened under normal conditions of heat treatment (i.e. as for commercial use). The experimental methods have proved quite adequate for this purpose and the results, in general, indicate that much useful information is likely to be forthcoming from the application of this technique to the study of age-hardening.

2. EXPERIMENTAL ARRANGEMENTS

(a) *Method*

The experimental methods have been described in full in prior publications (Sykes 1935; Sykes and Jones 1936). The principle is as follows: the specimen in the form of a closed hollow cylinder is placed inside and thermally insulated from a second closed cylinder of copper. The outer cylinder is heated by means of a resistance furnace, whilst the specimen is heated independently by means of a small auxiliary heating coil. By suitable adjustment of the heating currents the two cylinders can be held at about the same temperature and heat transfer consequently minimized. From the knowledge of energy input to the specimen and the resulting rate of rise of temperature, the instantaneous specific heat can be obtained as the temperature is continuously varied. The specific heat-temperature curve, subsequently referred to as the *ST* curve, consists of the normal specific heat-temperature curve called the *SCpT* curve with the thermal effects due to precipitation etc. superimposed. Since the *SCpT* curve in the absence of precipitation or allied phenomena can be obtained with reasonable certainty either by calculation or experimentally from the annealed specimen, the thermal effects due to precipitation can be obtained by difference. All the *ST* curves were obtained using a heating rate of about 2° C/min.; it varied from 1.6 to 2.1° C/min. as the temperature increased from 100 to 600° C.

(b) *Production of specimens*

Silver-copper alloys, 7.5 and 8.9 % Cu. These were prepared from high purity silver containing 99.9 % silver, and electrolytic copper containing 99.98 % copper. The alloying was carried out in a high-frequency furnace in a vacuum of about 0.05 mm. The ingots were forged and then homogenized at 770° C for 50 hr. *in vacuo*.

Aluminium-copper alloy, 4.8 % Cu. This was prepared from electrolytic copper and high purity aluminium (99.8 % aluminium). The melting was

carried out in air, and the alloy cast in a chill mould. The ingot was hammered out and then heat-treated at 525° C for 50 hr. in air.

Aluminium-magnesium-silicon alloy, 0.75 % Mg, 0.51 % Si. This was prepared from high purity aluminium, magnesium (99.98 % Mg) and silicon 98 % pure. The alloy was melted in air, and homogenized for 15 hr. at 525° C.

Copper-beryllium alloy, 2.6 % Be. This material in the form of bar 3 cm. diameter, was kindly supplied by Dr W. J. P. Rohn of the Heraeus-Vacuum-schmelze G.m.b.H., Hanau.

Duralumin. This was standard material kindly supplied by Mr S. J. Nightingale of Messrs James Booth and Co., Birmingham, and had the following nominal analysis (5 L. 1. Specification):

Cu	3.5–4.5 %	(4.1 %)
Si	< 0.7 %	(0.49 %)
Mg	0.4–0.7 %	(0.59 %)
Fe	< 0.7 %	(0.49 %)
Ti	< 0.3 %	—
Mn	0.4–0.7 %	(0.69 %)
Al	Remainder	—

The values in brackets are the results of analysis.

All the above materials were prepared in the form of bar, having a diameter of 1½ in.

X-ray powder photographs obtained from filings taken over the surface of the specific heat specimens indicated that no noticeable segregation was present in any of the samples.

(c) Hardness measurements

The hardness specimens were cut directly from the bars from which the specific heat specimens were machined without any additional forging treatment. Measurements were made using a Vickers machine, with a diamond pyramid and a 50 kg. load.

3. EXPERIMENTAL RESULTS

(a) Silver-copper alloy, 7.5 % Cu

Fig. 1 shows *ST* curves obtained on this alloy after various heat treatments. If the specimen is cooled at 80° C/hr. from 600° C practically all the copper is precipitated. The *ST* curve taken on heating is given in fig. 1*a* the circles being the experimental points. Using the values of *C_p* for silver

(Moser 1936) and copper (Jaeger and others 1932) we have calculated the specific heat of the alloy assuming it to be a pure mixture (Kopp-Neumann Rule). The results are plotted (X) in fig. 1*b*, giving the so-called ΣCpT curve, which should be a reasonable approximation to the true ST curve of the alloy when no change in atomic configuration is taking place.

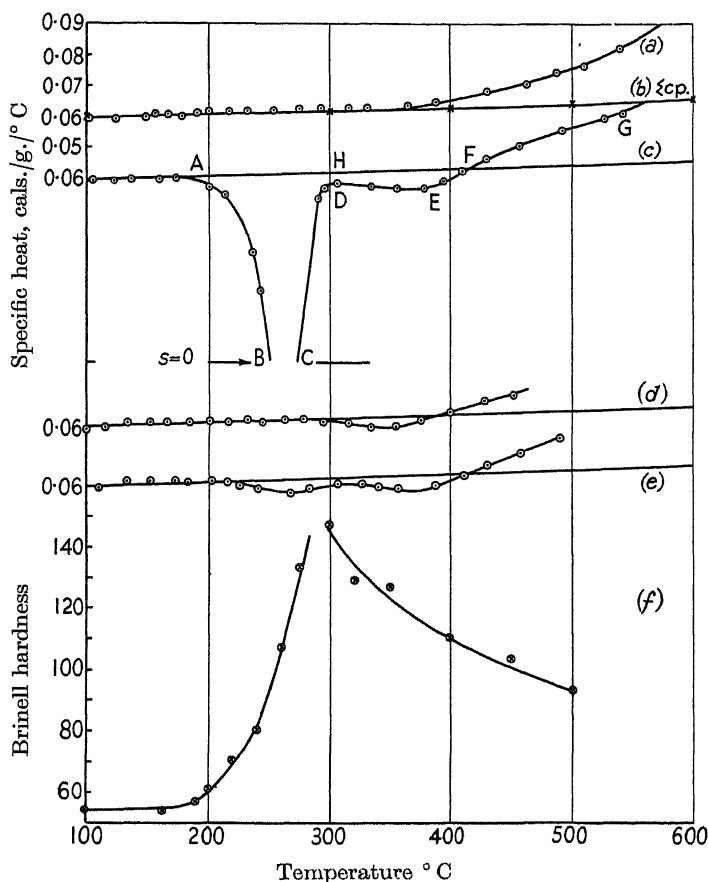


FIG. 1

Up to about 400° C the two curves are in agreement within the limits of experimental error $\pm 2\%$. Above 400° C the curves diverge, the apparent specific heat, S , of the alloy being higher than that calculated for the pure mixture, Cp . The difference is the "chemical energy" of heat of solution required to dissolve the precipitated copper which is gradually going back into solution.

Fig. 1c is the ST curve for the alloy as quenched from 760°C in water. Up to 180°C (A), it does not differ appreciably from 1a and 1b. The apparent specific heat then drops rapidly until at 250°C (B), it is zero. In this region the supersaturated solution is giving up its chemical energy. From 250 (B) to 280°C (C) the rate of heat evolution, at the heating rate employed, viz. 2°C/min. , is sufficient to heat the specimen without any additional external means, so that the apparent specific heat is either zero or negative. From 280 (C) to 300°C (D) the apparent specific heat rises very rapidly, then it remains fairly constant to 380°C (E) and finally rises gradually, crossing the ΣCpT curve at F .

We can associate the evolution of heat directly with precipitation; consequently it is reasonable to conclude that, under the experimental conditions used for fig. 1c, precipitation sets in at about 180°C , and gradually accelerates with increasing temperature along AB . The shape of the curve corresponding to diminishing rate of precipitation $CDEF$ is rather surprising, as it contains a marked point of inflexion at or about D ; only at F is the nett thermal effect, of the processes taking place in the alloy, zero. It appears therefore that the release of energy takes place in two stages; $ABCD$ due to precipitation and DEF . As the second phase which is precipitated in this alloy is substantially pure copper, it is unlikely that heat evolved in the stage DEF is produced by any reaction in the solid and we therefore associate it with the surface energy released by the gradual growth of the precipitated particles. (In a subsequent section an estimate of the size of the precipitated particles will be made from a consideration of the energy evolved along DEF .) It will be noted that 1c crosses the ΣCpT curve at a higher temperature and a greater angle than 1a. Re-solution should commence in 1c not later than 1a since the precipitate will be smaller. The apparent anomaly arises due to the fact that the ST curve 1c only shows the nett effect of the reactions in the specimen; i.e. re-solution is occurring along DEF , but is more than counterbalanced by heat evolution due to growth, at F the effects of the two processes are equal but opposite.

From 500°C upwards to 600°C the values of S obtained in 1a and 1c are the same within the limits of experimental error, indicating that solution is taking place at the same rate, and further that the precipitated particles are now so large that any energy evolved as a result of growth is negligible to a first approximation.

Fig. 1f is a hardness-temperature curve obtained as follows: a number of specimens were heat treated at 760°C and quenched in water. They were then heated at 2°C/min. and quenched at appropriate intervals and the hardness determined. The conditions are thus directly comparable with

those used during the determination of fig. 1*c*. Hardening commences at about 200° C, i.e. when a small amount of heat has been evolved—precipitation started and is a maximum when the second stage of heat evolution has just set in, viz. 300° C, and when the major portion of the chemical energy has been released, i.e. precipitation complete.

Fig. 1*d* is the ST curve of the alloy age-hardened at 300° C for 30 min. after quenching in water from 760° C. The Brinell hardness number was 126. The pronounced initial dip associated with precipitation is absent, but the particles are still small enough for their subsequent growth to release an appreciable amount of energy, i.e. the curve from 300 to 400° C is below the ΣCpT curve, in spite of any re-solution which may be taking place.

Fig. 1*e* is an ST curve of the alloy age-hardened at 200° C for 13 hr. after quenching in water from 760° C. The Brinell hardness number was 126. The result is very similar to fig. 1*d*, there being only a slight trace of the initial dip.

We conclude that high values of hardness produced by the normal ageing processes, i.e. in the temperature range 200–300° C, are typical of a state of the alloy in which the major portion of the chemical energy available in the original supersaturated solid solution has been released, i.e. precipitation has taken place.

Experiments carried out using a silver-copper alloy containing 8.9 % copper also gave results in complete agreement with those recorded in fig. 1.

(b) Copper-beryllium alloy, 2.6 % Be

Fig. 2 shows a number of ST curves obtained on this alloy. Fig. 2*a* refers to the alloy as cooled at 1° C/min. from 600° C. Up to 450° C the curve is identical within the limits of experimental error with the ΣCpT curve, fig. 2*b*, obtained by calculation using the specific heats of pure copper (Jaeger and others 1932), and pure beryllium (Jaeger and Rosenbohm 1936). Above 450° C the experimental curve diverges due to re-solution of the second phase.

Fig. 2*c* was obtained on the alloy as received (in the quenched state); the heat treatment given by Heraeus consisted in heating up to 750° C and quenching in water. The hardness was 160. The general behaviour is very similar to that recorded in fig. 1*c*. Evolution of heat sets in at about 160° C and is very rapid from 260 to 300° C; the second stage associated with the growth of the precipitated particles sets in at about 330° C and at 425° C the curve crosses the ΣCpT curve, and rises rapidly as the precipitated particles redissolve.

Comparison of fig. 2*c* with the hardness-temperature curve taken at 2° C/min., fig. 2*e*, indicates that maximum hardness, which occurs at a temperature of about 360° C, corresponds to a state of the alloy in which the energy in initial dip has been released.

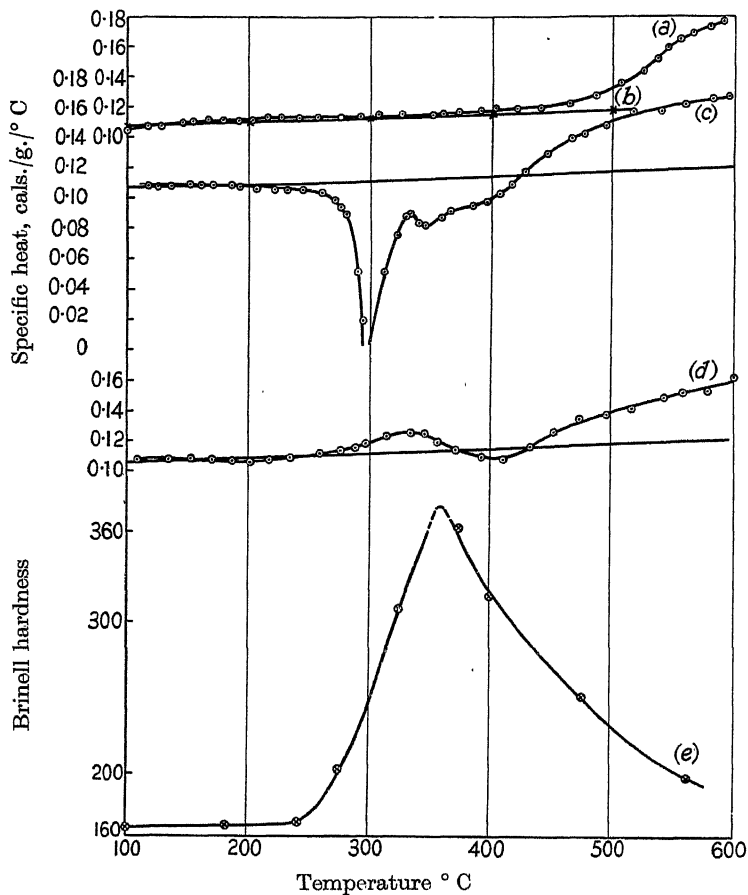


FIG. 2

Fig. 2*d* is the *ST* curve obtained on the material as quenched from 860° C and aged at 300° C for 150 min. and then cooled slowly to room temperature. The Brinell number was 370. This curve confirms the conclusion derived from figs. 2*c* and 2*e*, that maximum hardness is obtained after precipitation.

Whilst the copper-beryllium alloy and silver-copper alloy behave in a similar manner in that hardening proceeds in parallel with precipitation,

there are certain differences in the ST curves. The second dip is more pronounced in fig. 2*c* than in 1*c*, and in fig. 2*d* the ST curve rises above the ΣCpT curve before the heat, evolved as a result of the growth of the precipitated particles, becomes appreciable.

According to Hansen (1936), at temperatures above 576° C the β phase in equilibrium with the α , face centred solid solution of beryllium in copper is body centred and disordered. Below 576° C this phase decomposes into α and γ , γ being an ordered body centred phase of the approximate composition CuBe. None of the ST curves taken on the copper-beryllium alloy shows any discontinuity at 576° C, indicating that no transformation takes place in our material at this temperature, and there is therefore disagreement between our results and the equilibrium diagram published by Hansen. On extending the specific heat measurements to 700° C we have found a marked increase in specific heat at 626° C which corresponds, presumably, to the transformation temperature.

If the second phase undergoes a transformation in the solid state, then the ST curve obtained on heating the supersaturated solid solution is likely to be more complicated than that obtained for the silver-copper alloy in which the precipitated phase is substantially pure copper. The heat of solution of 1 g. atom of beryllium in copper is about 3600 cal.,* whereas the energy involved in disordering an equivalent amount of CuBe† is 900 cal. (which is of the same order). If the precipitated particles are ordered when released from the solid solution then we should expect the ST curves at temperatures up to 600° C to be very similar to those of the silver-copper alloy. However, in view of the probable small dimensions and high state of strain of the particles it seems likely that they will be precipitated in the disordered state, and at some stage in their subsequent growth ordering will take place and energy will be released which will augment that given out as a result of the reduction in surface energy of the particles. It is probable that the equilibrium degree of order may depend not only on the temperature but also on the size of the particles.

The experimental work is not yet sufficiently far advanced to enable an explanation to be put forward for the differences in the ST curves of silver-copper and copper-beryllium alloys, in particular the initial maximum in fig. 2*d*; for the present we wish to emphasize the similarity in behaviour namely that the hardened state is one in which the major portion of the chemical energy has been released, i.e. precipitation has taken place.

* By calculation from the variation of solubility with temperature.

† Corresponding to a critical temperature of 627° C.

(c) *Aluminium-copper alloy*

Fig. 3 gives a number of ST curves obtained on this material. Fig. 3a is the ST curve on the slowly cooled alloy which had the following treatment; 20 hr. at 370° C, 24 hr. at 310° C, 24 hr. at 270° C, 24 hr. at 230° C, 48 hr. at 160° C, 24 hr. at 110° C, and then cooled to room temperature at 30° C/hr. It is in reasonable agreement at low temperatures with the ΣCpT curve, fig. 3b. Above 300° C the curves diverge indicating that solution is taking place.

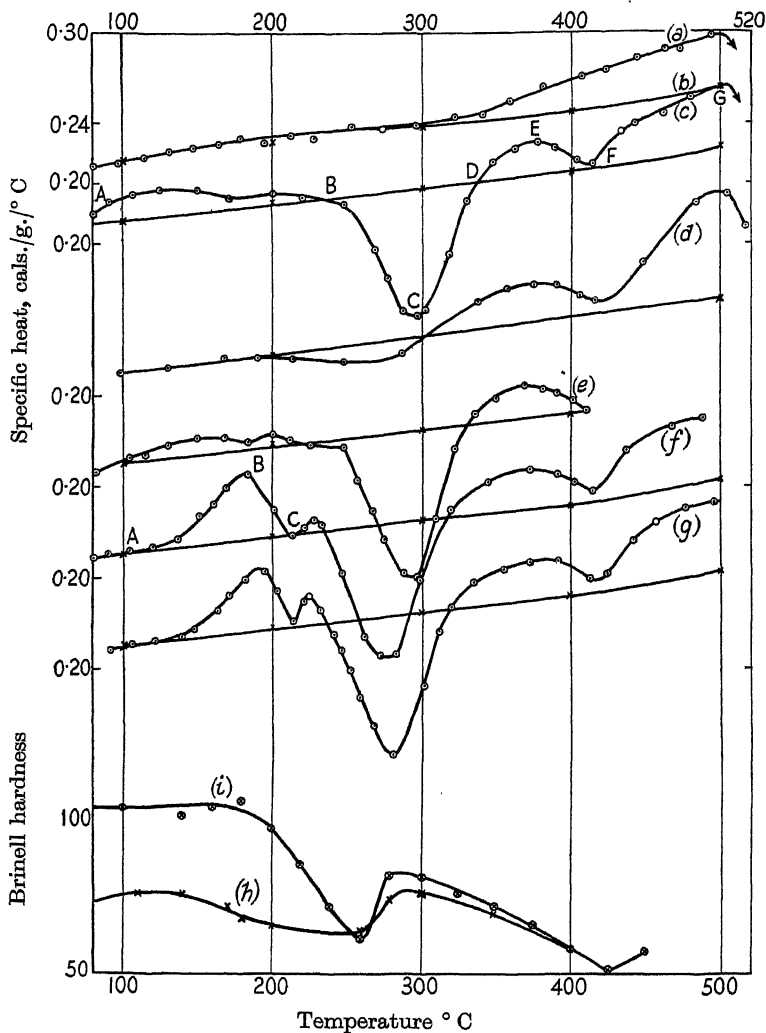


FIG. 3

Fig. 3*c* is the ST curve on the alloy as quenched from 525° C in cold water, and aged at room temperature for 32 days; the hardness was not measured on the specimen in this particular case, but experiments on material from the same cast show that hardening occurs at room temperature. The Brinell number would be greater than 80. This curve differs in several important features from fig. 1*c* and 2*c*. At temperatures from 100 to 220° C $A \rightarrow B$, the ST curve lies *above* the ΣCpT curve. The maximum difference is small (6 %), but it is greater than the experimental error (± 2 %), whilst the relative error which may affect the shape is not likely to be more than 0.5 %.

From 240° C onwards energy is evolved, rapidly attaining a maximum at about 300° C, indicating that precipitation is taking place, but the ST curve rises well above the ΣCpT curve, i.e. to E before any sign of a second change of heat evolution appears (cf. Figs. 1*c* and 2*c*).

Along $DEFG$ re-solution is taking place, as may be demonstrated from fig. 3*d*. In this case the specimen was quenched from 525° C into cold water reheated to 380° C, i.e. point E (fig. 3*c*), at 2° C/min., and then quenched. On reheating (fig. 3*d*), the material dissolved along DE (fig. 3*c*), has reprecipitated in the temperature range 200–300° C.

Reverting to fig. 3*c* from F onwards the apparent specific heat rises rapidly to G , and then diminishes approaching the ΣCpT curve again, indicating that complete solution, of the material precipitated along BC , has taken place. (Measurements have not been made above 520° C since, according to the equilibrium diagram, melting of the material is likely to take place. Experiments undertaken with an alloy containing 4 % Cu, where complete solution takes place below 520° C, confirm the above interpretation for the maximum at G .)

Fig. 3*e* is the ST curve for the alloy as quenched from 525° C and age-hardened at room temperature for 3 days. This curve is included to show that the same general features are present as in fig. 3*c*, and in particular that at temperatures up to 200° C the ST curve lies above the ΣCpT curve. Superposition of figs. 3*e* and 3*c* indicates that the excess energy supplied up to 200° C is higher in 3*c* than in 3*e*, a result which is attributed to the longer ageing period.

Fig. 3*h* is the hardness-temperature curve taken at 2° C/min. immediately after quenching from 525° C. The Brinell hardness of the material at the commencement of the experiment was 65. There is an increase in hardness up to about 140° C. In marked contrast with the results obtained with the silver-copper and copper-beryllium alloys softening occurs prior to and during precipitation, i.e. at 260° C; when precipitation is pro-

ceeding rapidly the hardness is 63, which is about the same as that of the freshly quenched material. The hardening accompanying precipitation is relatively small and softening commences before precipitation is complete.

This marked difference in behaviour is also evident in the ST curves of the hardened materials. Fig. 3*f* is the ST curve on the copper-aluminium alloy as quenched from 525° C aged at 110° C for 15 hr. The Brinell hardness number was 100. Fig. 3*g* is the ST curve on the alloy as quenched from 525° C and aged at 140° C for 22 hr. (Brinell hardness number 110.)

In both cases the ST curves are characterized by the fact that the first large dip associated with precipitation is still present (cf. figs. 1*d* and 2*d*), i.e. the hardening produced in the sample was not accompanied by precipitation. In both cases, also, there is no doubt that the ST curve lies well above the ΣCpT curve in the range 150–210° C; in fact, this is the noteworthy characteristic of the hardened alloy.

Fig. 3*i* is the hardness temperature curve taken at 2° C/min. on material water quenched from 525° C and aged at 140° C for 22 hr. (Brinell number 106.) This curve is directly comparable with 3*g*. The hardness decreases from 180° C onwards and the minimum occurs at 260° C, which corresponds to a temperature at which the rate of precipitation is almost a maximum. There is a small maximum in hardness at 280° C which could be associated with precipitation hardening, but softening sets in before precipitation is complete.

We may summarize the main results in fig. 3 as follows: high values of hardness can be produced by ageing at low temperatures, i.e. up to 200° C, without precipitation. The ST curve of the hardened alloy exhibits a marked maximum before the minimum associated with precipitation is reached; i.e. extra energy has to be supplied before the energy of precipitation can be released.

It will be noticed that the ST curves of the aluminium-copper alloy, fig. 3*c*, etc., are quite different from figs. 1*c* and 2*c* in that there is no second minimum immediately subsequent to the large initial dip attributed to precipitation. It is true that an additional minimum occurs at F after considerable re-solution has taken place, but the form of curve in this region is not what we should expect by comparison with figs. 1*c* and 2*c*. In this connexion it is also necessary to bear in mind that according to Wasserman and Weerts (1935) the precipitate from the supersaturated solid solution is not CuAl_2 but an intermediate phase. The minimum at F may result, due to the breakdown of this intermediate phase, into stable CuAl_2 . If such is the case, then the growth of the precipitated particles produces no noteworthy effects on the ST curve, and the particles as precipitated must

be very large. This would be in agreement with the experimental fact, fig. 3*i*, that the increase in hardness immediately subsequent to, or during, precipitation is relatively quite small.

Frenkel (1933) has taken a number of differential heating curves on a 4 % aluminium-copper alloy at a heating rate of 6–10° C/min. Comparison of these curves with those recorded in fig. 3 shows that whilst the positions of the maxima and minima differ to some extent, due probably to differences in material and technique, the same general shape is obtained. Frenkel concluded that no marked precipitation occurred during the normal ageing treatment given to this alloy.

(d) Aluminium-magnesium-silicon alloy

Fig. 4 gives a number of ST curves obtained on this material. Fig. 4*a* refers to the alloy slowly cooled, the annealing treatment being identical with that given to the aluminium-copper alloy (cf. fig. 3*a*). Fig. 4*b*, including the points marked \times , is the ST curve of pure aluminium, which should be a very near approximation to the ΣCpT curve for the alloy since the additions are very small. Comparison of figs. 4*a* and 4*b* suggests that re-solution commences at about 380° C. Fig. 4*c* is the ST curve of the alloy as quenched from 525° C and aged at room temperature for several days. Brinell hardness number 50. Up to 200° C, $A-B$, the curve whilst slightly higher is a very near approximation to the ΣCpT curve; energy is then released very rapidly, BC , and a second minimum appears at E , followed by a third minimum at G . (Two sets of points \times , \odot , are given on this curve. They refer to two different experiments using different specimens, and show that the results are reasonably reproducible.)

The second minimum at E occurs in the same place with respect to the first minimum as the second minimum in the curves for the silver-copper and copper-beryllium alloys (figs. 1*c* and 2*c*), and might therefore be associated with growth; consequently this alloy might be expected to harden appreciably on precipitation. (It is, of course, also possible that the magnesium and silicon precipitate separately, giving rise to two minima, or that the second minimum is due to combination of magnesium and silicon precipitated along BCD .) On the other hand, the fact that the ST curve rises above the ΣCpT curve and then falls very rapidly suggests that the material may harden prior to precipitation as in the case of the copper-aluminium alloy.

Fig. 4*f* is a hardness-temperature curve taken at 2° C/min. on material immediately after quenching from 525° C (Brinell hardness 29). There is a fairly rapid rise in hardness up to about 120° C, then a small drop. Once

precipitation sets in (200°C) the hardness rises rapidly to a maximum at 250°C , corresponding to the point *D* on the *ST* curve (fig. 4*c*).

Fig. 4*g* is a similar hardness curve on material aged at room temperature for several days and is comparable with fig. 4*c*. There is a slight drop in hardness before precipitation sets in, then a rise to a maximum at or about

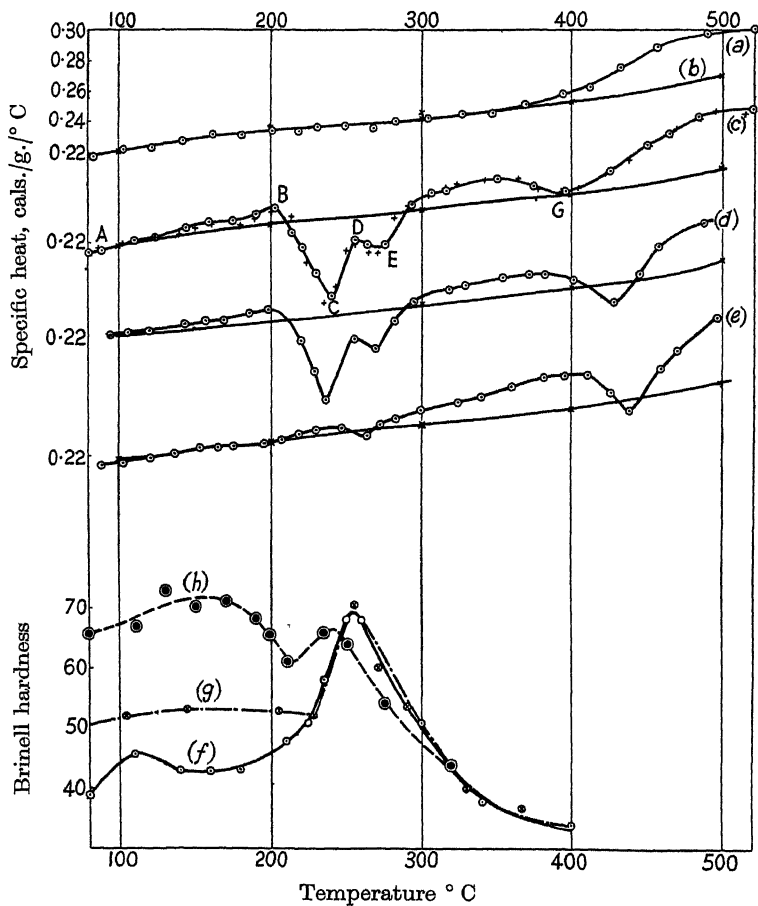


FIG. 4

260°C . It appears therefore that this material can either be hardened prior to precipitation like the aluminium-copper alloy, or after as is the case in the silver-copper alloy.

Fig. 4*d* is the *ST* curve of the alloy after ageing at 106°C for 16 hr. Fig. 4*h* is the hardness temperature curve obtained on material treated in the same way (Brinell number 67). Softening occurs prior to precipitation

and is followed by hardening on precipitation. There is, however, no marked maximum on the ST curve prior to precipitation (cf. figs. 3*f* and 3*g*).

Fig. 4*e* is the ST curve obtained on the alloy after ageing at 240° C for 100 min. (Brinell number 65). In this case although the material is hard all the energy in the first dip has been released by the ageing process and only the remains of the second dip are observed, i.e. the material is hard after precipitation has occurred. These results are therefore in agreement with those deduced from the ST curve (fig. 4*c*), and the hardness-temperature curve (fig. 4*d*).

(*e*) *Duralumin*

Fig. 5*a* is the ST curve of duralumin after cooling at 15° C/hr. from 480° C, and fig. 5*b* the ΣCpT curve calculated assuming it to be a pure mixture of its constituents. Fig. 5*c* is the ST curve after full ageing at room temperature (Brinell number 115). It is clear from the presence of the large initial dip that no appreciable precipitation has taken place during the hardening at room temperature. As in the case of the aluminium-copper alloy no well-marked second minimum is present and we might expect therefore that the amount of precipitation hardening would be relatively small.

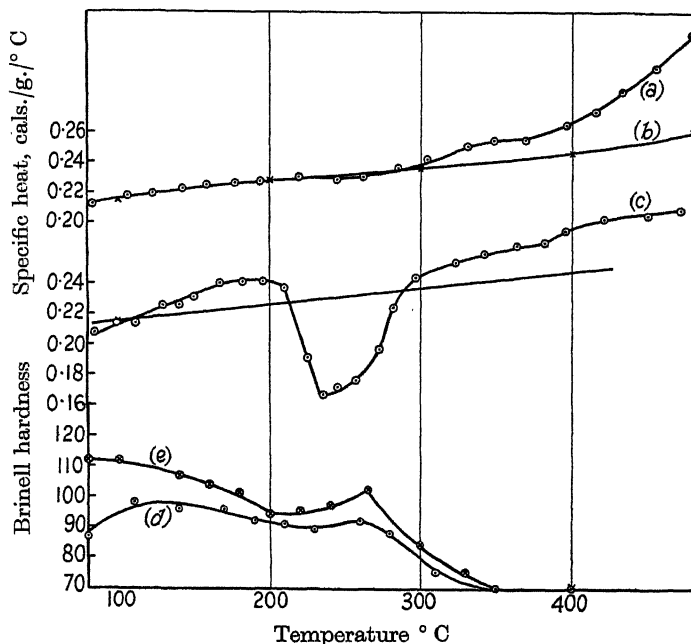


FIG. 5

Fig. 5*d* is the hardness-temperature curve taken at 2° C/min. on material immediately after quenching from 490° C. The increase in hardness during precipitation is small. Fig. 5*e* is a similar curve obtained on the material fully hardened at room temperature. As in the case of the aluminium-copper and aluminium-magnesium-silicon alloys, there is a drop in hardness before the increase due to precipitation hardening takes place. Softening, however, sets in before precipitation is complete, i.e. in this respect duralumin is similar to the aluminium-copper alloy and different from the aluminium-magnesium-silicon alloy.

4. HEATS OF SOLUTION

In the previous sections we have discussed in a qualitative manner a number of *ST* curves which refer to the precipitation of a second phase from a supersaturated solid solution. The heat of solution can be determined from such curves and in this section we shall compare the values so obtained with those calculated from the solid solubility curve using the well-known thermodynamic theorems.

Let us consider the equilibrium between a solid solution of two metals *A* and *B* and a second phase *C* composed of *A* and *B*, *A* being the major constituent of the solid solution. We shall consider two cases: one when the second phase is substantially pure *B*, i.e. the silver-copper alloy; and the other when the second phase is fully ordered, i.e. CuAl₂ in the aluminium-copper alloy. Let the concentration of the solid solution in equilibrium with the second phase be *c* at a given temperature *T*, the concentration *c* being expressed in gram atoms of solute per gram atom of the alloy.

Then the condition for equilibrium is that

$$Q = RT[\log c], \quad (1)$$

Q is the heat of solution, i.e. the heat required to dissolve 1 g.-atom of the second phase in an infinitely large amount of the solid solution of concentration *c*. It is equal to the change in internal energy occurring when 1 g.-atom of the second phase passes into solution. *Q* may vary with the concentration, unless the solution is very dilute.

The heat absorbed, *Hs*, when the concentration is increased from *c*₁ to *c*₂ is

$$Hs = R \int_{c_1}^{c_2} T[\log c] dc. \quad (2)$$

The relation between *T* and *c* is given by the solid solubility curve and *Hs* can therefore be evaluated using (2).

The derivation of equation (1) involves the assumption that the solid solution has an entirely random arrangement of atoms on the lattice sites at all temperatures.

Table I gives the solubility of copper in silver as a function of temperature. The values, taken from Hansen (1936), are the mean of the results of Stockdale, and Ageew and Sachs.

TABLE I. SOLUBILITY OF COPPER IN SILVER. Wt. %

Temp. °C	779	750	700	600	500	400	300
Solubility %	8.8	7.2	5.5	3.3	1.8	1.0	0.65

The solubility of silver in copper is less than 1 % at 400° C, so that the deposited phase is practically pure copper.

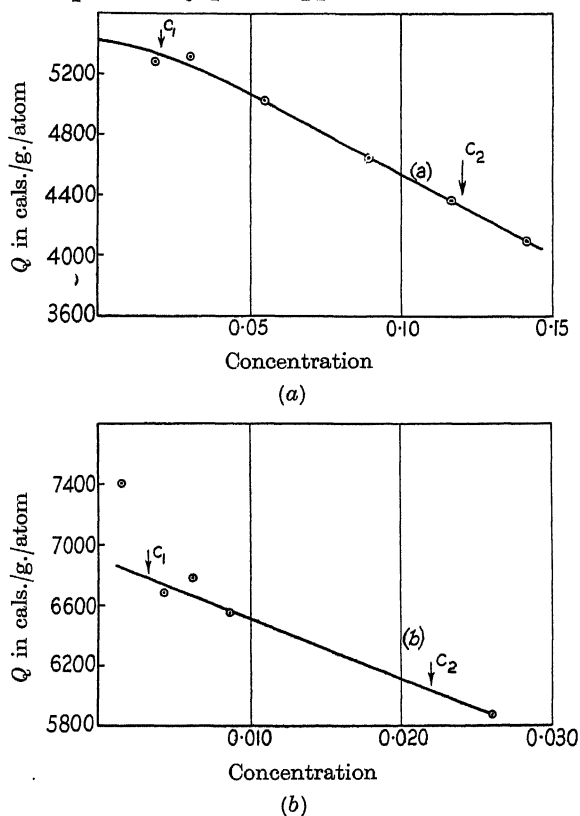


FIG. 6

Using the above experimental data $Q = RT[\log c]$ has been plotted as a function of the concentration c . (c is the number of gram atoms of copper dissolved in 1 g.-atom of the solid solution.) Fig. 6a: it is seen that Q

diminishes as c increases, as one would expect. Before H_s can be evaluated it is necessary to decide on the limits for the integration. Our specimens were quenched from 760°C at which temperature silver holds 7.5 % copper in solution, $c = 0.12$. This concentration is therefore the upper limit. The lower limit has also to be chosen with due regard to the experimental procedure if reliable comparison with the ST curves is to be made. The lower limit was fixed at 400°C as this is the temperature at which resolution becomes appreciable on the annealed material (cf. fig. 1*a*). Considerable variation in the lower limit does not affect the magnitude of H_s appreciably as the change in solubility with temperature is small at 400°C . H_s was therefore evaluated over the range $400\text{--}750^\circ\text{C}$ and the value so obtained is 4.70 cal./g.

The solid solubility curve of the aluminium-copper alloys was taken from Hansen (1936) (see Table II).

TABLE II. SOLUBILITY OF COPPER IN ALUMINIUM. Wt. %

Temp. $^\circ\text{C}$	548	430	400	350	300
Solubility %	5.65	2.0	1.4	0.95	0.34

The identity of the phase deposited from the aluminium-copper solid solution in our experiments is not very certain. For the purposes of the estimation we have assumed it to be CuAl_2 (the concentrations are so small that it does not make any material difference to the answer what constitution is assumed for the second phase provided its molecular weight is of the same order as CuAl_2). The particular alloy used contained 4.8 % Cu, $c = 0.022$, corresponding to complete solution at 525°C , so that these values were taken for the upper limit of the integration. The lower limit was taken at 300°C . The values of Q as a function of c are shown in fig. 6*b*, there is a considerable amount of dispersion amongst the points indicating that the values are not too reliable. The value of H_s between the limits specified is 4.24 cal./g.

Experimental determination of H_s . Copper-silver alloy

If the specimen is quenched from temperature T to temperature T_0 as a homogeneous solid solution and no changes take place in the atomic configuration during quenching then the total heat given out is

$$H = \int_T^{T_0} Cp dT. \quad (3)$$

If the specimen is then reheated from T_0 to T the energy absorbed must also

be equal to H . If, on the other hand, the annealed specimen is heated from T_0 to T the energy required is

$$= \int_{T_0}^T C_p dT + Hs. \quad (4)$$

Consequently Hs is given by the difference in the energy required to heat the quenched and annealed specimens over the temperature range T_0 to T , the experiment being carried out in such a way that both specimens pass through temperature T in equilibrium.

Calorimetric measurements at temperatures up to 800° C have not yet attained sufficient precision to enable Hs to be determined accurately in this way. The value of H for the silver-copper alloy from 100 to 800° C is 45 cal./g. and for the aluminium-copper alloy from 100 to 525° C about 100 cal./g., so that an error in H of 1 % produces an error in Hs of from 25 to 50 %. (The errors in H arise chiefly in the temperature range from 500° C upwards where difficulties are encountered due to lack of reproducibility in thermocouples. Our measurements are limited to a maximum temperature of 600° C.) There are, however, methods of arriving at a fairly reliable value of Hs from the ST curves.

In the case of the silver-copper alloy the ST curves of the annealed and quenched specimens agree very well over the temperature range from 520° C upwards, indicating that at 520° C the specimens are in the same condition. The difference in energy from room temperature to 520° C can therefore be put equal to Hs . A number of experiments on the alloy containing 8.9 % Cu gave a difference of 4.4 cal./g. The value obtained for the 7.5 % Cu alloy was 4.42 cal./g. The probable error is ± 4 %.

The evaluation of Hs for the aluminium-copper alloy cannot be carried out satisfactorily by the above method for two reasons: (a) the specimens change appreciably after quenching and before the measurements can be made: (b) the total energies required to heat the specimens from 100 to 525° C is very large compared with the heat of dilution. A good approximation can be obtained by integrating the energy between the ST curve and the $C_p T$ curve over the temperature range in which precipitation takes place, i.e. B to D (fig. 3c). This has been done for the curves 3c, 3e, 3f, 3g. The values found vary from 4.10 to 4.40 cal./g. (This variation of ± 0.15 cal./g. is less than 1 % of the *total* energy absorbed over the temperature interval, i.e. it is of the same order as the experimental error. Within these limits the heat evolved during precipitation is the same irrespective of the hardening treatment.)

If the material precipitated from the solid solution is some intermediate

phase as isolated by Wasserman and Weerts (1935) and not CuAl_2 then especially good agreement with the theoretical result is not to be expected. It is conceivable that the minimum at F is associated with the formation of CuAl_2 from this intermediate phase and this additional energy of 1.1 cal./g. should be added to the experimental value of 4.2 cal./g., giving 5.3 cal./g.

The experimental values of H_s obtained for the two alloys are in reasonably satisfactory agreement with those calculated theoretically, indicating that variations in apparent specific heat attributed to precipitation have been correctly chosen. Similar comparisons have not yet been made for the other alloys used in this investigation.

5. THE ST CURVE FOR THE SILVER-COPPER ALLOY, FIG. 1c

We have suggested that the precipitation process can be split up into two parts; one giving rise to the large evolution of heat represented by the area $ABCDHA$, and the other which we have ascribed to the growth of the particles and represented by $HDEFH$ (neglecting any effects due to resolution).

We might expect the precipitation process to proceed somewhat as follows: first, segregation of copper atoms will take place at certain selected positions on the silver lattice, this will diminish the entropy and heat will be evolved. Secondly, when sufficient copper atoms have collected together they will break away from the silver lattice and take up a lattice spacing roughly equal to that of normal copper. This will reduce the strain energy in both copper and silver and heat will be evolved. There will still be considerable strain in the vicinity of the boundaries of the precipitated particles, which will be released as growth proceeds giving rise to further heat evolution, this third stage sets in at D . Whilst the first two stages may proceed simultaneously to a great extent it is of interest to point out that the ST curve shows no trace of any discontinuity which could be attributed to the inception of the second stage and it is doubtful whether any details as to the precise mechanism of precipitation proper, i.e. the second stage, are likely to be forthcoming from specific heat data. (It might be argued that the area $HDEFD$ represents the heat given up when the copper atoms break away from the silver lattice. This is very unlikely since in the case of the aluminium-copper alloy we should arrive at the conclusion that precipitation did not take place in our experiments until a temperature of 380° C was attained. This is not in agreement with other experimental work on this alloy.)

In order to make an estimate of the size of the particles corresponding to the state of maximum hardness, point *D*, fig. 1*c*, we shall assume that the particles are equal in size and cubic in shape. Then provided the linear dimensions are reasonably large compared with one atomic distance, the central portion of the particle will take up the normal lattice spacing of copper and the strain will be confined to the boundary atoms. We shall assume that all the strain is located in the surface layer of atoms and that these take up some intermediate spacing between that of copper and silver.

If *d* is the linear dimension of the particles in atomic distances and *N* their number per gram of silver-copper alloy, then the number of copper atoms deposited from 1 g. of the alloy is

$$Nd^3 = 0.06 \times 6 \times 10^{23} / 63.6$$

(0.06 g. of copper being precipitated per gram of alloy).

The number of atoms in the surface of the particles is

$$6Nd^2.$$

These atoms will fit neither the silver lattice nor the copper lattice and may be described as being in an amorphous state since they possess no specific crystalline pattern. The extra energy per gram atom will be of the same order as the latent heat of melting, viz. 2700 cal.

Consequently the strain energy per gram of silver-copper alloy is

$$6Nd^2 \times 2700 / 6 \times 10^{23}.$$

This, according to our interpretation of the *ST* curves, is equal to the energy represented by the difference in area of the curves, figs. 1*c* and 1*a* from 300° C upwards. This is 0.7 cal./g. i.e.

$$\frac{6Nd^2 \times 2700}{6 \times 10^{23}} = 0.7.$$

These two relations involving *N* and *d*, enable *N*, the number of particles to be eliminated and *d* is found to be about 20 atomic distances.

The above assumptions are, of course, a very crude approximation. When the copper is precipitated the lattice dimensions of the silver will change and slip may occur; some energy will therefore be stored in the silver lattice. If the strain energy is not restricted to the surface of the particles the estimated value of *d*, viz. 20 atomic distances, will be too small.

Furthermore, it is unlikely that the disturbance at the boundary of the particle is restricted to a monatomic layer and this would mean that the magnitude of *d* had been underestimated. Similarly the precipitation of

thin plate-like particles would give a larger value for the average volume of the particles.

It appears, therefore, that a linear dimension of 20 atomic distances, viz. 75 Å, is a lower limit of size for the precipitated particles corresponding to maximum hardness provided the interpretation given to figs. 1*a* and 1*c* is the correct one.

6. THE EFFECT OF AGEING ON THE ST CURVE OF THE ALUMINIUM-COPPER ALLOY

As mentioned in § 3 (*d*) the ST curve of the hardened alloy is characterized by the fact that it rises above the ΣCpT curve in the temperature range 150–220° C.

The variation in apparent specific heat with temperature over and above that expected in accordance with the ΣCpT curve is produced by changes in the atomic configuration of the specimen; and the areas enclosed between the ST curve and the ΣCpT curve are a measure of the energy absorbed or released by these changes, e.g. the area $BCDB$ (fig. 3*c*), is a measure of the energy evolved during precipitation. From 140 to 210° C (fig. 3*f*), energy is *absorbed* when the hardened alloy is heated through this range of temperature, and it is of interest to consider what changes in the atomic arrangement of the solid solution are responsible for this absorption of energy.

The material as hardened is in a metastable state, the equilibrium state being that in which the major portion of the copper is precipitated as CuAl_2 , consequently the free energy, F , of the hardened alloy at room temperature is not a minimum. Any changes which take place in the material will tend to reduce the free energy.* The relation between the free energy, the internal energy U and the entropy S is given by

$$F = U - TS, \quad (5)$$

where T is the absolute temperature. As the temperature dependent parts of F , U and S are taken care of by the CpT curve, we shall consider that the various parameters in (5) are affected only by changes in the state of the alloy.

Referring to fig. 3*f* we see that on heating from room temperature the internal energy U is increasing since the apparent specific heat s is greater than Cp ; consequently since F cannot increase the entropy S must also be increasing. The maximum entropy of a solid solution is obtained when the

* We consider the free energy instead of the thermodynamic potential since the volume changes are very small.

solute atoms are arranged at random; it follows therefore that in the hardened alloy the copper atoms are not arranged at random, i.e. some segregation has taken place. In the alloy as quenched it is highly probable that a random arrangement exists, the hardening process is therefore accompanied by a rearrangement of the copper atoms into a state of lower entropy. Such rearrangement will give rise to the evolution of heat and this has been observed in the case of the age-hardening of duralumin at room temperature. Swietoslawski and Czochralski (1936) reported a heat evolution of 0.47 cal./g. during ageing for duralumin quenched from 510° C. Czochralski and others (1937) have recently improved the technique of measurement of such small heat evolutions and find a value of 1.67 cal./g. The heat evolved during ageing should be equal to or greater than the extra energy put in before precipitation occurs. According to fig. 5c this extra energy is 1.4 cal./g., which is in fair agreement with the findings of the Polish workers.

The nature of the "ordered" or hardened solid solution cannot be deduced from the ST curves; two dimensional segregation along planes (Desch 1934) or the formation of clusters of copper atoms in the solid solution would lead to the required diminution in entropy.

It can be argued that some rearrangement of the copper atoms must take place on the aluminium lattice before precipitation can occur, e.g. in the case of the silver-copper alloy clustering of copper atoms must occur on the silver lattice prior to precipitation if it is conceded that the copper cannot be precipitated in the form of single atoms but as particles of finite size. The interesting point in connexion with the copper aluminium alloy is that some or all of these clusters *disperse* again before precipitation sets in. If this were not the case then the ST curve for the hardened copper-aluminium alloy would always be below the ΣCpT curve (cf. fig. 1c) for the silver-copper alloy, until complete precipitation had been effected. Perhaps a little speculation on this point may be permitted. Let us consider the process of precipitation and assume that a stable particle cannot be precipitated unless it contains a finite number of atoms. The chance of such a particle forming depends amongst other factors on the rate of atomic interchange and the number of atoms of solute *in solution* which can take part in the interchange. In the case of the aluminium-copper alloy the precipitated phase has a complicated structure and it is conceivable that the smallest particle which can be deposited is relatively large, say, compared with those deposited from the silver-copper alloy. The chance of such a particle forming and precipitating is therefore very small, and an intermediate stage of free energy lower than that of the random solid solution yet higher than that corre-

sponding to true equilibrium may be reached which will remain quite stable at low temperatures. This state corresponds to the hardened state, in which some segregation has occurred, and in order to picture the processes occurring on heating we shall assume that the copper atoms have grouped themselves into a large number of small clusters which stiffen up, i.e. harden the aluminium solid solution. These clusters will be surrounded by material which contains only a very small amount of copper corresponding to a first approximation to the solid solution saturated at room temperature. When the temperature is increased the solid solubility also increases so that some copper atoms will be absorbed by the solid solution until it is saturated again, i.e. softening will occur. At the same time the large clusters will also tend to grow at the expense of the small ones, and this will give rise to hardening. The former process involves movement of copper atoms over distances appreciably smaller than the latter so that one might expect that when a piece of aluminium-copper alloy hardened at a low temperature is heated it would first soften and then harden. No claim is made that the mechanism postulated above is entirely responsible for the well-known phenomenon of "dehardening" of duralumin, etc.; the general argument is given to show that specific heat measurements may lead to a better understanding of the phenomenon.

As mentioned in a previous section of the paper, the energy evolved during precipitation represented by the area $BCDB$ (fig. 3c), is not affected appreciably by short period ageing (cf. figs. 3f and 3g); further, the hardness-temperature curves show a minimum hardness at about 260° C which is practically the same as that of the freshly quenched material. Consequently it appears likely that clusters disperse almost completely before precipitation sets in. If this conclusion is correct then the energy evolved during ageing should be equal to that represented by the area enclosed by the ΣCpT curve and the ST curve prior to precipitation, i.e. $ABCA$ (fig. 3f). Suitable experiments will enable this to be checked. At low temperatures the clusters do not disperse completely, provided precipitation does not occur, and consequently the material does not soften completely before re-hardening.

7. GENERAL

G. I. Taylor and his collaborators (Quinney and Taylor 1937) have shown that the energy bound up with lattice distortion in highly cold-worked metals is of the order of 0.2–0.5 cal./g. The experimental results on the age-hardening alloys used in this investigation show that the chemical

energy available in the supersaturated solid solutions is about 5 cal./g., i.e. roughly ten times that which can be stored in the material as a result of lattice distortion. Consequently the "transfer" efficiency in age-hardening need not be particularly high. Any mechanism which enables some of the chemical energy to be stored in the form of lattice strains will produce hardening. Previous experimental work has given rise to the view that there are at least two processes which produce strains during age-hardening; one by the formation of knots or clusters of solute atoms in the parent lattice, and the other by the formation of a large number of finely divided particles of the second phase. Gayler has proposed, and hardness measurements tend to confirm, that both processes may take place in the same material at constant temperature, the second process following the first.

The experimental work described in the present paper shows that in the case of the silver-copper alloy maximum hardness is associated with a state of the alloy in which precipitation has taken place, i.e. the second process is operative. The hardness-temperature curve shows no maximum which could be attributed to knot formation. This is not surprising in view of the very small effects found by Cohen. On the other hand, the *ST* curves on the aluminium-copper alloy prove conclusively that knot formation takes place, and is accompanied by appreciable hardening on ageing at low temperatures. The increase in hardness on precipitation is however relatively small. The experiments on the aluminium-magnesium-silicon alloy and duralumin confirm the conclusion that hardening can take place both prior to and after precipitation, although much further work will be necessary to determine to what extent and at which stage each particular type of solute atom contributes to the hardening.

It appears, therefore, that Gayler's view that both processes occur during age-hardening is essentially correct. The present work, however, provides no definite proof that both processes always occur in a given alloy at *the same temperature* the one following the other. Gayler's theory is based primarily on ageing experiments carried out at constant temperature, whereas our experiments are made at the temperature increases at a constant rate, and this difference must be considered in correlating the results. For example, whilst it is clear that the aluminium-magnesium-silicon alloy can be hardened at 100° C, i.e. by the first process, and also at 240° C by the second, it is quite likely that the second process will not set in at 100° C in a finite time, and it seems highly probable that the room temperature hardening of the aluminium alloys is due entirely to the first process. At higher temperatures prolonged heat treatment should lead to precipitation

and hardening by the second process will take place. This will necessitate a diminution in the number of knots or clusters in the solid solution and the nett effect may be a softening.

The statement by Gayler that softening should set in once precipitation proper starts is not in agreement* with the experimental results on the silver-copper alloy, since in this case maximum hardness occurs only when precipitation is complete.

The maximum values of hardness obtained on the silver-copper and copper-beryllium alloys are produced by precipitation hardening. In the aluminium-copper alloy and duralumin hardening by cluster formation is certainly as effective if not more effective than hardening on precipitation. The experimental work at present available is not sufficient to enable any reliable conclusions to be drawn as to the reasons for this variation in behaviour. Various factors suggest themselves and some have been discussed briefly in §§ 4 and 6, but further experimental work is necessary before the effects of these factors can be correctly ascertained.

The strain energy stored in the alloys in condition of maximum hardness is quite small (cf. figs. 1c and 1d), and specific heat measurements of higher precision than those recorded in this paper will be necessary to establish the relation between stored energy and hardness. In the case of alloys which harden by cluster formation it is not clear how the strain energy can be determined by specific heat measurements alone, since the heat evolved or absorbed is the difference between the change in internal energy produced by the atomic rearrangement and the strain energy set up by this rearrangement. Some independent method of arriving at the change in entropy will have to be developed before the strain energy can be determined from the ST curves.

In the preceding sections we have dealt only with the features which could be interpreted with reasonable certainty, and have been primarily concerned in demonstrating that maximum hardness may occur either before or after precipitation. A detailed study of the effect of a wide range of heat treatments on the ST curve of each particular alloy suitably correlated with measurements of other physical properties, such as crystalline structure, mechanical strength, electrical conductivity, etc., is desirable before a reliable and complete interpretation of the ST curves recorded in

* It is possible that the meaning we attribute to precipitation is different from that implied by Gayler. Precipitation is complete, according to our view, when the majority of the copper atoms have taken up their correct spacing, but the particles are still connected to the silver lattice at the boundaries. If by precipitation Gayler visualized particles free from strain then our results do not conflict with her views.

this paper can be given. For this reason a number of peculiar features in the *ST* curves of the various alloys have received little or no discussion.

ACKNOWLEDGEMENTS

We are indebted to Mr P. Leech, B.Sc., for his expert assistance in the determination of the many specific heat measurements recorded in this paper.

One of us (N. S.) is indebted to the Department of Scientific and Industrial Research for financial assistance, the other (C.S.) is a member of the Research Panel responsible to the British Iron and Steel Federation for research on the structure of alloys.

The work has been carried out partly in the Metallurgical Department of the University of Cambridge, under the direction of Professor R. S. Hutton, M.A., D.Sc., and partly in the Research Department of the Metropolitan-Vickers Electrical Co., Ltd., Trafford Park, Manchester, under the direction of Dr A. P. M. Fleming, C.B.E.

We thank Professor W. L. Bragg, F.R.S., Professor R. Peierls and Dr D. Stockdale, for their kind interest in the investigation.

SUMMARY

The changes in atomic configuration taking place during hardening affect the apparent specific heat of age-hardening alloys. Specific heat-temperature curves and hardness measurements have been obtained on five typical age-hardening alloys.

The results show that in certain cases, e.g. the silver-copper and copper-beryllium alloys, maximum hardness is attained when the major portion of the chemical energy associated with the supersaturated solid solution has been evolved, i.e. precipitation has taken place. In other cases, duralumin and aluminium-copper alloys, maximum hardness is associated with segregation of the solute atoms in the parent lattice prior to precipitation.

REFERENCES

- Bragg and Williams 1935 *Proc. Roy. Soc. A*, **152**, 231.
Bridgman 1931 "Physics of High Pressure". Edinburgh.
Cohen 1936 "Metals Technology", p. 751.
Czochralski and others 1937 *Wiad. Inst. Met. Warsaw*, p. 45.
Desch 1934 "Chemistry of Solids". Ithaca.
Frenkel 1933 *Metallwirtsch. Metalwiss.* **12**, 583.

- Gayler 1937 *J. Inst. Met.* **60**, 1, 249.
 Gayler and Preston 1932 *J. Inst. Met.* **48**, 197.
 Hansen 1936 "Aufbau der Zweistoff Legierungen". Berlin.
 Jaeger and others 1932 *Proc. Acad. Sci. Amst.* **35**, 772.
 Jaeger and Rosenbohm 1934 *Proc. Acad. Sci. Amst.* **37**, 67.
 Merica, Waltenburg and Scott 1919 *Bur. Stand. J. Res., Wash.*, no. 347.
 Moser 1936 *Phys. Z.* **37**, 737.
 Mott and Jones 1936 "Properties of Metals and Alloys", p. 23. Oxford.
 Quinney and Taylor 1937 *Proc. Roy. Soc. A*, **163**, 157.
 Swietoslawski and Czochralski 1936 *Wiad. Inst. Met. Warsaw*, p. 60.
 Sykes 1935 *Proc. Roy. Soc. A*, **148**, 422.
 Sykes and Jones 1936 *J. Inst. Met.* **59**, 257.
 Wasserman and Weerts 1935 *Metallwirtsch. Metallwiss.* **14**, 605.

The classical equation of state of gaseous helium, neon and argon

BY R. A. BUCKINGHAM, PH.D.

Queen's University, Belfast

(Communicated by J. E. Lennard-Jones, F.R.S.—Received 11 July 1938)

The deviations from the equation of state for perfect gases which are observed in all known gases result from the interactions of their constituent atoms or molecules. The excess pressures observed at all but the lowest temperatures show that the dominating factor is the strong repulsion between atoms at close range, due to the interpenetration of complete electron shells. Little is known about these repulsions, and that is readily summarized. Between atoms with spherically symmetrical distributions it is likely that the repulsive potential is accurately represented by a function

$$P(r)e^{-r/\rho}, \quad (1)$$

where r is separation of the atomic nuclei and $P(r)$ a polynomial in r . Quantum theoretical calculations made by Slater (1928) for helium atoms (with a closed shell of two electrons) and by Bleick and Mayer (1934) for neon atoms (with a closed shell of eight electrons) show that an adequate expression may sometimes be obtained if the polynomial is replaced by a constant. Some confirmation of this (though over a very restricted range of r) is given by Born and Mayer (1932) and Huggins (1937), whose work

on ionic cubic crystals shows that their elastic properties are admirably correlated when the repulsive potential of two ions of rare gas type is represented by an exponential function* $be^{-r/\rho}$, with a range of about one atomic diameter.

Between neutral atoms the more important interaction at larger distances is the attractive van der Waals' force arising from the mutual polarization of the atoms. The van der Waals' energy can be represented according to quantum theory by a series

$$-(cr^{-6} + dr^{-8} + gr^{-10} + \dots), \quad (2)$$

in which c , d , etc., are constants; although at distances of two atomic diameters or more the only important term is that in r^{-6} . The expression (2) is usually assumed to be valid outside a sphere equal in volume to the polarizability of the atoms concerned. Now there is an intermediate range of r , which includes the zero and minimum of the interatomic potential and throughout which the penetration and polarizability effects are of comparable importance. It is not readily possible in this region to predict the form of the potential energy function, for it is probable that the two kinds of interaction are not even approximately independent and in any event the expansion (2) is here of doubtful validity. Since evidence by direct calculation can be obtained only with great labour, and is restricted by the accuracy of available atomic wave functions, any information which can be deduced from experimental data is welcome. For this purpose the study of gases is more fruitful than that of solids, yielding information over a wider range of r , or more precisely, it enables the restricted information obtainable from the properties of solids to be extended to a wider range. In the first part of this paper we shall follow the method of approach of Lennard-Jones (1924) to this subject, and then introduce some further developments.

THE SECOND VIRIAL COEFFICIENT

The equation of state of a gas may be written as

$$pV = NkT \left\{ 1 + \frac{B}{V} + O\left(\frac{1}{V^2}\right) \right\},$$

where V is the volume of a gram-molecule, and B is a function of temperature

* However, a general study of these properties has led J. A. Wasastjerna (1932) to conclude that at the distances concerned the dominant term is that in $r^7e^{-r/\rho}$.

and of the intermolecular energy. For a spherically symmetrical field $E(r)$, the classical formula for B is

$$B = 2\pi N \int_0^\infty r^2 (1 - e^{-E(r)/kT}) dr, \quad (3)$$

when V is measured in gram-molecular units at N.T.P.*

Now quantum effects are likely to be important only at low temperatures and for light atoms such as helium, as shown for example by Uhlenbeck and Beth (1936). If we wish not merely to obtain an empirical expression for $E(r)$ which gives good agreement between the observed values of B and those calculated from (3) over the widest possible range of temperature, but rather to learn something of the true interaction, it is important to exclude those temperatures for which the classical formula is certainly invalid. For this purpose it is permissible to use the classical formula, even for helium, provided that as little weight as possible is attached to low temperature observations. Thus for helium we shall neglect experimental data below 70° K, though indeed quantum deviations are appreciable even at 300° K and it is necessary to discuss the consequences of their omission.

If the potential $E(r)$ is known, it is a simple matter to evaluate the virial coefficient from (3). However, this paper is concerned with the problem of deducing the form of $E(r)$ from the observed virial coefficients, and no progress is possible unless a fairly simple form is adopted for $E(r)$. One which has often been used is

$$E(r) = \lambda r^{-s} - \mu r^{-6}, \quad (4)$$

in which s is a constant greater than 6, and the positive term is intended to represent the short range repulsive field discussed above. If λ and μ are slowly varying functions of r this function is sufficiently flexible to represent the true interaction accurately over the whole range. To assume this at once however leads to great analytical difficulties, and as a first approximation it is necessary to assume that λ and μ are both constant, in which case the problem is analytically soluble. It then appears that (4) is an adequate representation over a small range of r only, and so as a second approximation we can increase the flexibility of the function by allowing λ to vary slowly (μ still being regarded as constant). At the same time, the formulae derived on the assumption that λ is constant may still be taken as reliable over a short range of r . We thus obtain more precise knowledge of

* Note that the numerical value of B given by (3) is 4.46×10^{-5} times the value obtained when V is measured in c.c.

the repulsive potential. This is a rough outline of the procedure we shall adopt.

As a beginning then, λ and μ are assumed constant. The function (4) then has a similar asymptotic behaviour to the van der Waals' energy (2), but it does not follow that the constant μ that we shall derive from the virial coefficient is comparable with the true van der Waals' constant c . It may be expected to be greater than c , for two reasons. The main contribution to the virial coefficient made by the negative term in (4) arises from the region in which dr^{-s} is appreciable; the omission of the r^{-8} term will therefore tend to be compensated for by an increase in the coefficient of r^{-6} . Secondly, the inverse power function λr^{-s} does not fall off as rapidly as one containing an exponential factor. It may happen that if s and λ are such that the variation of $E(r)$ when positive is fairly well represented, then the contribution of $E(r)$ in the middle region covering the zero and minimum of $E(r)$ is unduly large. This also can be redressed by an increase in μ beyond the value which it should otherwise have, say beyond the minimum. On these grounds, the parameter μ must be regarded as no less empirical than s or λ .

With λ and μ as constants, Lennard-Jones (1924) derived a closed formula for B , which is as follows:

$$B = \frac{2}{3}N\pi(\lambda/\mu)^{3/s-t}F(y), \quad (5)$$

where
$$F(y) = y^{3/s-t} \left\{ \Gamma\left(\frac{s-3}{s}\right) - \sum_{n=1}^{\infty} c_n y^n \right\}, \quad (6)$$

and y is a function of temperature given by

$$y = \frac{\mu}{kT} \left(\frac{kT}{\lambda} \right)^{t/s}. \quad (7)$$

The coefficients c_n are given by

$$c_n = 3\Gamma\left(\frac{nt-3}{s}\right) / n! s. \quad (8)$$

These equations are subject only to the conditions $s > 3$, $t > 3$, which are here fulfilled, since t represents the power of r in the negative term, set equal to 6 in (4). From them Lennard-Jones was able to derive values of the parameters λ and μ which for various integral values of s gave good agreement between the observed virial coefficients and those of equation (5), over a wide range of temperature. His method of doing this, which we need not describe in detail,* was simple and convenient and entailed only the

* It is described in the original paper by Lennard-Jones (1924) and by Fowler (1936).

superposition of the theoretical and experimental curves of $\log B$ against $\log T$. The chief disadvantage is that the observations at lower temperatures, and especially those below the Boyle point temperature, are given an importance which, in view of remarks above about the validity of the classical formula, is quite contrary to what we desire. This criticism, of course, applies mainly to helium, which has a Boyle point at about 23°K ; for the heavier gases, it matters little. However, the difficulty can be overcome by using a different method to derive the parameters λ, μ which does not depend upon the inclusion of low temperature observations of B .

As a first approximation we aim at finding average values of λ and μ (for given s) which lead to good agreement over a fairly wide range of temperature. Let new quantities X and Y be defined by

$$X = (\lambda/\mu)^{3/(s-t)}, \quad Y = (\mu^s/\lambda^t)^{1/(s-t)},$$

and equations (5) and (7) rewritten in the form

$$\log X = \log \frac{3}{2\pi N} B - \log F(y), \quad (9)$$

$$\log Y = \frac{s}{s-t} \log y + \log kT. \quad (10)$$

By definition $\log X$ and $\log Y$ are also related to $\log \lambda$ and $\log \mu$ by the equations

$$\log \lambda = \frac{1}{3}s \log X + \log Y, \quad (11)$$

$$\log \mu = \frac{1}{3}t \log X + \log Y. \quad (12)$$

For a given temperature B can be taken from the observations and its value inserted in (9). Equations (9) and (10) then determine a relation between $\log X$ and $\log Y$ in parametric form (the parameter being y), since $F(y)$ is a known function, s and t being given. This relation can be shown as a curve in the $(\log X, \log Y)$ plane. The intersection of two such curves for observed values of B at two temperatures determines $\log X$ and $\log Y$, and therefore $\log \lambda$ and $\log \mu$, uniquely. In practice one requires a series of values of B at regular temperature intervals. If the form $\lambda r^{-s} - \mu r^{-t}$ for the chosen values of s and t were an adequate representation of $E(r)$, and if the values of B contained no experimental errors, all the curves would pass through a point; actually most of the intersections will lie within a fairly well-defined area. By taking the median values of the ordinates and abscissae of all the intersections one gets suitable average values of $\log X$ and $\log Y$, and thence by equations (11) and (12) $\log \lambda$ and $\log \mu$.

Turning now to the application of this procedure to the rare gases, we make use of the extensive measurements made by Holborn and Otto (1924-6) of the virial coefficients of helium, neon and argon. Some of their results are given in Table I, and included also are a few values for helium (marked *N*) due to Nijhoff, Iliin and Keesom (1928).

TABLE I. OBSERVED VALUES OF THE SECOND VIRIAL COEFFICIENT

Helium		Neon		Argon	
$T^{\circ}\text{K}$	$B \times 10^3$	$T^{\circ}\text{K}$	$B \times 10^3$	$T^{\circ}\text{K}$	$B \times 10^3$
71.5 N	0.458	65.1	-0.935	173	-2.871
89.9 N	0.481	90.5	-0.365	223	-1.687
90.0	0.473	173	0.288	273	-0.986
126.5 N	0.553	223	0.406	323	-0.492
169.7 N	0.589	273	0.475	373	-0.192
273	0.529	373	0.529	423	0.052
323	0.523	473	0.582	473	0.208
373	0.510	573	0.614	573	0.501
473	0.493	673	0.612	673	0.683
573	0.469				
673	0.454				

These have been used in the manner described to obtain the results shown in Table II, which contains λ and μ for five integral values of s between 8 and 14 (t having the value 6 throughout).^{*} For any of the three gases considered, all the forms derived for $E(r)$ lead to fairly satisfactory values of the virial coefficients over the temperature ranges covered by Table I, showing that the experimental data do not determine, except within wide limits, the most suitable value of the parameter s . Moreover, as s increases from 8 to 14 the form of $E(r)$ given by (4) changes markedly, and it is clear that little physical significance can be attached to the evidence afforded by virial coefficients alone about the atomic interactions, at least as far as the classical theory is concerned.

This difficulty led Lennard-Jones (1931) to consider the rare gases in their solid crystalline forms, as the lattice constants and energies of these crystals depend upon the depth and position of the minimum of the interaction energy $E(r)$. With additional evidence about this point it is possible to pick out from Table II the energy function for each gas which has most physical importance as follows.

^{*} Except for helium, the results in Table II do not differ much from those published by Lennard-Jones (1931), although the values of λ and μ vary more regularly with s as a result of the different method of derivation. For comparison Lennard-Jones' values for $s = 12$ are included in the table. The author is greatly indebted to Professor Lennard-Jones for the use of many of his original calculations.

TABLE II. AVERAGE VALUES OF λ AND μ DERIVED FROM THE VIRIAL COEFFICIENTS

s	Helium		Neon		Argon	
	$\lambda \times 10^{8(s+1)}$	$\mu \times 10^{60}$	$\lambda \times 10^{8(s+1)}$	$\mu \times 10^{60}$	$\lambda \times 10^{8(s+1)}$	$\mu \times 10^{60}$
8	0.0030 ₄	3.70	0.0183	20.7	0.315	237
9	0.0051 ₀	2.37	0.0350	14.5	0.768	170
10	0.0094 ₈	1.73	0.0732	11.4	2.05	137
12	0.0356	1.04	0.355	8.32	16.2	103
14	0.139	0.76	1.82	6.78	136.5	86.7
L.-J. (1931)						
12	0.0380	1.24	0.36	8.45	17.7	108.3

Value of λ in erg. cm.⁸, and of μ in erg. cm.⁶

Neon and argon are known to crystallize in the face-centred cubic form. Let us consider a unit cell containing four atoms in a face-centred cubic lattice in which the shortest distance between atoms is a . We will assume that the interaction of any atom with its six nearest neighbours is represented by (4), but that for more distant neighbours the asymptotic form of the field is valid, i.e.

$$E(r) \sim -(cr^{-6} + dr^{-8}).$$

The potential energy *per unit cell* at 0° K is then

$$\phi(a) = \frac{24\lambda}{a^8} - \frac{24c}{a^6} - \frac{2(C_6 - 12)c}{a^6} - \frac{2(C_8 - 12)d}{a^8}, \quad (13)$$

where C_6 and C_8 are constants evaluated by Lennard-Jones and Ingham (1925) and having the values

$$C_6 = 14.454, \quad C_8 = 12.802.$$

There is a departure here from the practice of Lennard-Jones in that (4) is not assumed to lead to the correct asymptotic form of the field. Now the equilibrium value of a , say a_0 , corresponds to the minimum of ϕ , whence it follows that

$$\frac{s\lambda}{a_0^{s-6}} = 6\mu + 1.227c + \frac{0.534d}{a_0^2}, \quad (14)$$

and the corresponding energy of a crystal containing N atoms is

$$U_0 = \frac{1}{4}N\phi(a_0). \quad (15)$$

Thus given s , λ , μ , c and d equations (14) and (15) determine the values of a_0 and U_0 . The van der Waals' constants for the rare gases have been discussed

elsewhere by the author (Buckingham 1937*a, b*), and it is believed that the values given in the following table are of the right order of magnitude:

TABLE III. VAN DER WAALS' CONSTANTS

	Helium	Neon	Argon
<i>c</i>	1.28	6.1	60×10^{-60} erg cm. ⁶
<i>d</i>	2.75	8.8	180×10^{-76} erg cm. ⁸

The contents of this table and of Table II for neon and argon have been applied to give Table IV, which contains the calculated values of a_0 and U_0 for the different values of s and also the observed values for 0° K estimated by extrapolation. The observed values of $-U_0$ include not only the observed heat of evaporation (extrapolated to 0° K) but also the zero-point vibrational energy (calculated from $\frac{3}{8}Nk\Theta_D$), since it is the sum of these quantities which must be compared with the crystal energy calculated from (15).

TABLE IV. VALUES OF a_0 (IN Å) AND U_0 (IN CAL./MOLE)
FOR NEON AND ARGON

	Neon		Argon	
<i>s</i>	a_0	$-U_0$	a_0	$-U_0$
8	3.33	351	4.09	1150
9	3.22	412	3.98	1340
10	3.13	470	3.88	1520
12	3.00	569	3.75	1820
14	2.92	655	3.66	2075
Exp.	(3.20)	590	3.81	2030

The experimental values of U_0 are taken from Clusius (1929) for Ne, F. Born (1922) for A. The value of a_0 for Ne may be slightly less than 3.20 Å, which is the measurement of de Smedt, Keesom and Mooy (1930) at about 4.2° K. Some writers have used 3.04 Å at 0° K, but this seems too small. The value 3.81 Å for argon is extrapolated from 3.82 Å at 21° K (de Smedt and Keesom 1926) and 3.83 Å at 40° K (Simon and v. Simson 1924).

It appears from Table IV that it is not possible to choose s to give good agreement between the calculated and observed values of the lattice constant and also of the crystal energy. Thus the values for $s = 9$ for Ne and $s = 11$ for A are in fair agreement with the measurements of a_0 , but equally good agreement for U_0 is only possible if s is rather greater than 12 for neon and rather less than 14 for argon. This contradiction may arise because the function (4) is unable to represent the interatomic energy over a range so wide as the properties of the solid *and* gaseous forms of the elements require.

However, it seems preferable to use the crystal energies to determine s

as they are more sensitive to changes in s . The calculated values of a_0 are then too small, the deficit for neon (for $s = 12.5$) being about 7 % or slightly less, and for argon ($s = 13.6$) about $3\frac{1}{2}$ %. It is interesting also to compare the corresponding values of μ , as given by Table II, with the van der Waals' constants. One should compare μ with the "effective" van der Waals' constant defined by

$$c' = c + da_0^{-2},$$

which allows for the magnitude of the r^{-8} term in the neighbourhood of the minimum of the crystal energy function. For neon, μ has the value 7.7 compared with 7.0 for c' ; whilst for argon the respective numbers are 90 and 75. So that in both cases μ is the larger, and the expectation of this result is seen to be justified. Indeed, but for the fact that we have chosen s as large as is consistent with the crystal data, the differences would be even greater. Owing to the lack of data for crystalline helium* it is not possible to determine the best value of s for helium with any precision, but it will be seen later that 12 is very nearly correct.

The conclusion reached so far is that the comparison of the observed virial coefficients with the theoretical formula (3), using the form (4) for $E(r)$, does not lead to an energy function which is even an approximate representation of the actual potential energy of two atoms, although with the help of crystal data we can derive a function which may have such significance in the middle range (including the zero and minimum of the energy) which makes the decisive contribution to the integral (3). This function however has not the correct asymptotic form for large r , nor in all probability does it represent adequately the asymptotic behaviour of the interaction for small r as given by (1). These discrepancies arise because the energy represented by (4), with μ and λ both constant, is too inflexible. Accordingly we introduce the second approximation in which λ , but not μ , is allowed to vary slowly with r . We notice that if the virial coefficient at any one temperature is known, if μ and s are given particular values, then λ is uniquely determined by equations (5), (6) and (7), and can be estimated in the following simple way. When $\log X$ and $\log Y$ in equation (12) are replaced by (9) and (10) respectively, one obtains

$$\log \mu = 2 \log \frac{3B}{2\pi N} + \log kT + \frac{s}{s-6} \log y - 2 \log F(y) \quad (t = 6)$$

* However, recent measurements by Keesom and Taconis (1938) suggest that the solid helium has a closely packed hexagonal structure, which may provide a useful basis for detailed calculation.

which may be rewritten as

$$\frac{s}{s-6} \log y - 2 \log F(y) = \log \left(\frac{3B}{2\pi N} \right)^2 \frac{\mu}{kT}. \quad (16)$$

Now the left-hand side of this equation is a function of y and s only and for a given value of s can be plotted without difficulty. The right-hand side is a numerical constant, since μ , B and T are assumed known; hence the value of y which satisfies the equation can be read at once from the graph of the left-hand side. It follows also from (7) that

$$\log \lambda = \frac{s}{6} (\log \mu - \log y) - \frac{s-6}{6} \log kT, \quad (17)$$

and so when y has been obtained as above, $\log \lambda$ can be evaluated.

The question arises whether this value of λ has any significance. It has been assumed that λ is a slowly varying function of r , whereas equation (5) was derived on the assumption that λ is constant. Now if we examine the classical formula for B , it appears that the integral can conveniently be divided into two parts according as r is less or greater than r_0 , where $r = r_0$ is the zero of $E(r)$. When $r < r_0$, $E(r)$ is positive and so is the integrand, and when $r > r_0$, $E(r)$ and the integrand are both negative. At temperatures below the Boyle point therefore, when B is negative, the second half of the integral is the larger, and one can state roughly that the attractive part of the field is the more important in determining the virial coefficient. Conversely at temperatures above the Boyle point the repulsive part of the field is relatively more important than the attractive. Thus at different temperatures different parts of the atomic field are effective in causing the deviations from perfect gas behaviour and, when we consider a particular temperature as above, the value of λ deduced is that which gives a good representation of the potential field *over the range in which B is most sensitive to changes in the field*. This region may be expected to lie near to but outside the classical distance of closest approach; on either side of it B is much less sensitive to inaccuracies in $E(r)$, so that in the first approximation λ may be assumed to have the same value at all points. Nevertheless, because the critical range depends upon the temperature, moving inwards as the temperature rises, we may find that λ is a function of temperature and in fact this is generally the case. Thus if s and μ are taken from Table II, the values of λ tend to increase in most cases with temperature, showing that the repulsive potential is more accurately represented by a term λr^{-s} in which λ is a decreasing function of r .

The next step is to investigate this variation of λ with r . Let us return to consider the results at a particular temperature. We will suppose that the mutual potential of two atoms is accurately represented, except at large distances, by

$$E(r) = b(r) - \mu r^{-6}, \quad (18)$$

in which $b(r)$ is the repulsive potential. Then when λ is determined from (17) for given s , μ and T , the expression λr^{-s} is an approximation to $b(r)$ valid over a limited range of r characteristic of the temperature concerned. Within certain limits however there is no reason for preferring one value of s to another, and the analysis may be repeated with a second value, say s' , leading to an expression $\lambda' r^{-s'}$ which is another approximate form for $b(r)$ valid over the same range of r . It is then not unreasonable to expect that the intersection of the two curves λr^{-s} and $\lambda' r^{-s'}$ lies within the range and is a point on or near the true curve $b(r)$. If r is equal to x at this point, then

$$\lambda x^{-s} = \lambda' x^{-s'}$$

$$\text{or} \quad (s' - s) \log x = \log \lambda' - \log \lambda, \quad (19)$$

and the corresponding value of the energy, by the above hypothesis,

$$E(x) = \lambda x^{-s} - \mu x^{-6}. \quad (20)$$

Now we have seen that x and $E(x)$ are both functions of temperature; consequently if the procedure is carried out for different temperatures a series of points should be obtained which are on or near the true energy curve, and these would provide an idea of the form of this curve, at least within certain limits of r . Calculations on these lines have therefore been made for helium, neon and argon, and we shall proceed to describe the results.

HELIUM

The experimental data of Holborn and Otto, Nijhoff and others (Table I) are found to give information about the interaction of helium atoms in a range 2.2–2.6 Å, which lies just within the zero of the energy function. Two facts emerged in the course of the calculations: first, that the curve for $E(x)$ given by (20) is almost independent of the choice of s and s' , as it must be if the method is to be useful. It is not evident from (19) that the intersection x does not depend on this choice, but actually the variation is small. Thus calculations were made for $s = 9$ and for two values of s' viz. 8 and 12, and it appeared that $\log \lambda'/\lambda$, considered as a function of s' , varies almost linearly. Secondly, the dependence of the curve for $E(x)$ upon the

choice of μ is more marked, but the variation when μ is allowed to vary within reasonable limits is not too great to prevent some definite conclusions about the true energy curve. A description of the calculations will indicate the extent of this variation. The following values were chosen for μ : 1.63×10^{-60} , which is the upper limit to the van der Waals' constant c , derived from the polarizability of helium (Buckingham 1937*a*), and also the "effective" van der Waals' constant likely to prevail near the minimum of the energy function; 1.47×10^{-60} , which is the value adopted by Slater (1928); and a smaller value 1.35×10^{-60} cm.⁶. For each of these three values a series of points was obtained by applying equations (16)–(20). These did not lie on smooth curves, owing to experimental errors in the virial coefficients. However, it was found that the points obtained by plotting $\log b(x)$ as a function of x are quite adequately represented by a straight line, so it was assumed that

$$b(r) = be^{-r/\rho}. \quad (21)$$

The most probable values of the constants b and ρ were then derived by the usual statistical formulae for linear regression. These are shown in Table V, together with the constants for the repulsive potential deduced by Slater (1928) using quantum mechanics. It will be seen that for the value of μ chosen by Slater, the agreement for ρ is extremely good, but the value of b is somewhat greater than Slater's.

TABLE V. REPULSIVE POTENTIAL CONSTANTS FOR HELIUM

$\mu \times 10^{60}$	$b \times 10^{10}$ (ergs)	ρ in Å	$2r_0$ in Å
1.35	12.6	0.208	1.48 ₃
1.47	8.71	0.216	1.46 ₂
1.63	5.06	0.229	1.42 ₇
Slater			
1.47	7.7	0.217	1.44 ₃

Note. The quantity $2r_0$ in this table is defined by $b(r) = e^{(2r_0-r)/\rho} \times 10^{-12}$ erg, which is the form used by Born and Mayer (1932).

Another aspect of the results is shown in fig. 1A. The actual points obtained for $E(r)$ using $\mu = 1.47$ are there shown together with the smooth curves for $\mu = 1.35$ and 1.63 calculated from the constants in Table V. Some of the low temperature points lie outside these limits, and generally speaking the experimental results between 70 and 200° K seem rather unreliable. Nijhoff's result for 170° K is particularly bad and was rejected in the calculations for b and ρ . However, for our purpose this matter is not

important as the classical formula is itself less reliable at temperatures below 200° K. The main point is that between 2.2 and 2.4 Å, which excludes these lower temperatures, the curve for $E(r)$ is fairly well defined and indicates a somewhat larger repulsive potential than is contained in the S field. In fig. 1B the field corresponding to $\mu = 1.47$ (curve B) is shown beside Slater's.

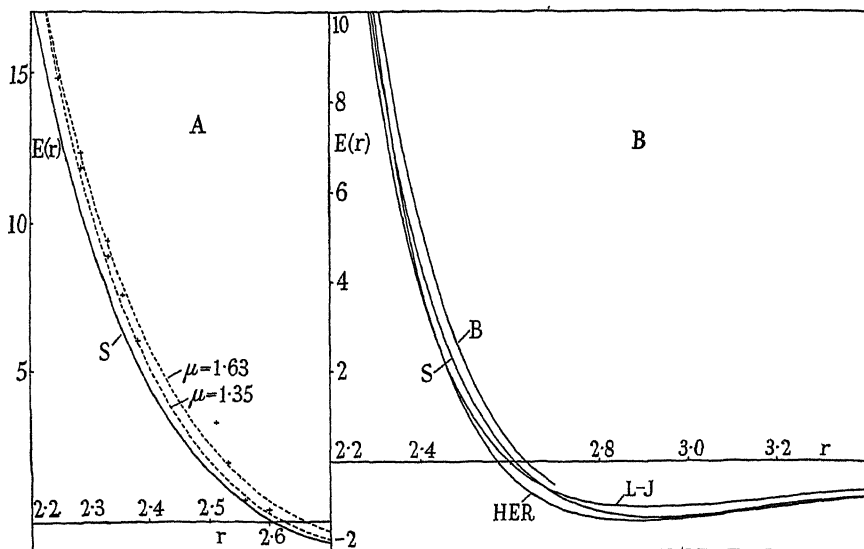


FIG. 1. Interaction of helium atoms. A. Interaction at small distances. Crosses + represent points obtained from experimental data with $\mu = 1.47 \times 10^{-60}$. B. Comparison of the following potential functions:

$$\left. \begin{array}{l} S. \text{—Slater} \\ B. \text{—Buckingham} \end{array} \right\} E(r) = be^{-r/\rho} - \mu r^{-6}, \text{ with } \mu = 1.47 \times 10^{-60}.$$

$$\left. \begin{array}{l} L. \text{—J.—Lennard-Jones} \\ H. \text{—E. R.—Hirschfelder, Ewell and Roebuck} \end{array} \right\} E(r) = \lambda r^{-s} - \mu r^{-6}, \text{ with } s = 12.$$

$E(r)$ is measured in units of 10^{-15} erg and r in Å.

In discussing the available evidence for the true energy function two points must be remembered. First, the direct calculations of the second virial coefficient by Kirkwood and Keyes (1931), using the classical formula and the Slater field, led to results less than Holborn and Otto's observed values, the deficit being 5 % at 250° K and 7 % at 350° K. This in itself suggests that the S field underestimates the repulsive potential at short distances. But secondly, it is necessary to take account of the quantum correction to the classical formula, which is almost certainly positive above 100° K. As this correction is not known with any certainty, especially

below normal temperature, it has not been allowed for in our calculations, though it is desirable to do so. If we accept as roughly correct the estimates made by Uhlenbeck and Beth (1936), the correction is about $3\frac{1}{2}\%$ at 250 and $2\frac{1}{2}\%$ at 350° K, so that at these temperature the Slater field values are too low only by $1\frac{1}{2}$ and $4\frac{1}{2}\%$ respectively, although at higher temperatures the difference may increase more rapidly. On the other hand, the fields deduced in these papers may be expected to give results which are too positive by an amount equal to the quantum correction. If 2.4 Å is taken to be the distance of greatest importance at 250° K it seems likely that the true interaction energy in this neighbourhood lies nearer to the *S* field than to field *B* (fig. 1B). At shorter distances it approaches nearer to *B*, but beyond 2.4 Å it probably falls below the *S* field at a point within the energy zero.*

Turning now to the negative part of the interaction, it is certain that as a consequence of the classical formula being assumed accurate, the helium fields in Tables II and V underestimate considerably the attractive potential between 2.6 and 3.0 Å. Moreover, the evidence obtained from the calculations at very low temperatures of the viscosity and second virial coefficient, discussed in the accompanying paper, is that at greater distances the *S* field has an attractive potential which is too large. We believe that the asymptotic potential is more nearly given by

$$E(r) \sim -1.28 \times 10^{-60} r^{-6}, \quad (22)$$

the constant being that calculated by Baber and Hassé (1937) using Hylleraas wave functions; it is 87 % of the coefficient of r^{-6} in the *S* field. Near the minimum of the field the difference is more than balanced by the r^{-8} term, which has a coefficient probably not greater than 2.75×10^{-78} erg cm.⁻⁸, a likely value being about 2.0×10^{-76} . The result is that the minimum is displaced inwards and is slightly deeper.

This seems to give a fairly good picture of what the true interaction should look like, on the assumption that the repulsive potential can be represented by a simple exponential. It is interesting to see how it compares with other functions using the inverse power form, and for this purpose two other functions have been drawn in fig. 1B, both involving r^{-12} . One of these (marked L.-J.), viz.

$$E(r) = 3.80 \times 10^{-10} r^{-12} - 1.24 \times 10^{-12} r^{-6}, \quad (r \text{ in Å}) \quad (23)$$

* This argument is based upon the Holborn and Otto data. The situation is more complicated if one takes into account the observations of Gibby, Tanner and Masson (1929), who obtained virial coefficients between 300 and 450° K, 3 % less on the average than those of H. and O. These would require the true energy to be nearer to the *S* function.

was given by Lennard-Jones in 1931 as one of several functions which agreed equally well with the gas data, for at that time the many calculations since carried out bearing on the properties of helium were not available to indicate how any particular function might be in error. As regards (23), the potential barrier is slightly lower than in either of the fields (*S* and *B*) which use exponentials, and the negative potential outside the zero is also somewhat less. The asymptotic form for large *r* is however almost identical with (22). Actually between the rough limits 2.6 and 3.5 Å the attractive field is smaller than it should be, because it was not possible then to take account of the quantum correction. The effect of this can be seen from the curve marked H.E.R., which represents

$$E(r) = 4.39 \times 10^{-10} r^{-12} - 1.522 \times 10^{-12} r^{-6}. \quad (24)$$

This function was given recently by Hirschfelder, Ewell and Roebuck (1938) and is based on experimental Joule-Thomson coefficients, allowance being made for quantum effects by assuming the Uhlenbeck and Beth corrections to be accurate. The increased attraction is very marked. If in the H.E.R. function an exponential were used in place of r^{-12} , one would expect a small increase in the potential barrier, combined with a minimum slightly deeper and sharper.

TABLE VI. REPULSIVE POTENTIAL CONSTANTS AND CALCULATED CRYSTAL PROPERTIES FOR NEON AND ARGON

	$\mu \times 10^{60}$	$b \times 10^{10}$	ρ	$2r_0$	a_0	U_0
Neon	7	745	0.175	1.98 ₀	2.92	740
	10	10.85	0.259	1.80 ₀	3.11	550
	12	3.90	0.292	1.74 ₂	3.19	490
				Obs.	(3.20)	590
Argon	100	224	0.266	2.66 ₈	3.74	2050
	120	26.0	0.328	2.58 ₂	3.85	1740
	150	4.18	0.408	2.46 ₅	3.93	1430
				Obs.	3.81	2030

NEON AND ARGON

Similar calculations have been completed for neon and argon, using Holborn and Otto's results in Table I. Various values of μ were assumed, and the appropriate parameters for the repulsive potential, expressed as a simple exponential, were derived in the same way as for helium. The results are shown in Table VI. For neon, the energy function corresponding to $\mu = 10 \times 10^{-60}$ is based upon a series of points between 2.6 and 3.0 Å and

covering energies between $+10$ and -4×10^{-15} erg; for argon with $\mu = 100 \times 10^{-60}$, the corresponding ranges are $3.4\text{--}3.6$ Å and -4.5 to -15.5×10^{-15} erg. For other values of μ similar but slightly different limits apply.

The energy functions depend very much upon the value of μ , and the properties of the crystalline elements have therefore been considered once more to select the most probable form for each interaction. The values of the lattice constant and crystal energy in Table VI have been calculated from the energy per unit cell:

$$\phi = 24be^{-r/\rho} - \frac{4}{r^6} \left\{ 6\mu + 1.227c + \frac{0.534d}{r^2} \right\}, \quad (25)$$

assuming as before that the interaction (18) applies only to nearest neighbours. The crystal energy is particularly sensitive to changes in μ , so we may assume that the most probable value of μ , and hence the most probable energy function, is that which gives a crystal energy of the order of the observed value. Thus for neon, the field for which $\mu = 9 \times 10^{-60}$ erg cm.⁶

$$E(r) = 25.7 \times 10^{-10} e^{-r/0.235} - 9 \times 10^{-12} r^{-6} \text{ ergs} \quad (r \text{ in Å}) \quad (26)$$

(the constants being obtained by interpolation), gives $U_0 = 600$ cal./mole and $a_0 = 3.05$ Å, and may be regarded as giving the best correlation of the properties of gaseous and crystalline neon, so far as these are known at present. It must be remembered, however, that this function is intended only as an approximate form for the interaction between 2.6 and 3.1 Å and that outside these limits its use is extrapolatory. It is illustrated (curve I) in fig. 2A and compared with the function previously obtained with a repulsive potential λr^{-12} (curve II, which may equally well be assumed to represent the function given by Lennard-Jones (1931)). The similarity between the two curves is of course partly due to the fact that they have been chosen to give roughly the same properties to the crystalline neon, but it also vindicates the use which Lennard-Jones made of the simple form (4) as a first approximation. One observes that the function with the exponential involves a slightly greater repulsion at small distances, balanced by an increased attraction further out.

With regard to neon we should mention the repulsive potential derived by Bleick and Mayer (1934) from *a priori* calculations. They calculated the potential at three points, 1.8 , 2.3 and 3.2 Å, and fitted the results by an exponential $be^{-r/\rho}$, in which $b = 190 \times 10^{-10}$ erg and $\rho = 0.209$ Å. The value of ρ agrees reasonably well with that in (26). At 2.3 Å the repulsive potential is two and a half times as great as that given by (26), though at 3.2 Å it

is only 30 % greater. Bleick and Mayer used wave functions for neon which disregard electron exchange and tend to overemphasize the charge distribution on the boundary of the atom. It is possible that the use of Hartree-Fock functions would reduce the calculated potential in this region.

For argon, the most satisfactory function, judging by Table VI, seems to be

$$E(r) = 1.69 \times 10^{-8} e^{-r/0.273} - 102 \times 10^{-12} r^{-6} \text{ ergs}, \quad (27)$$

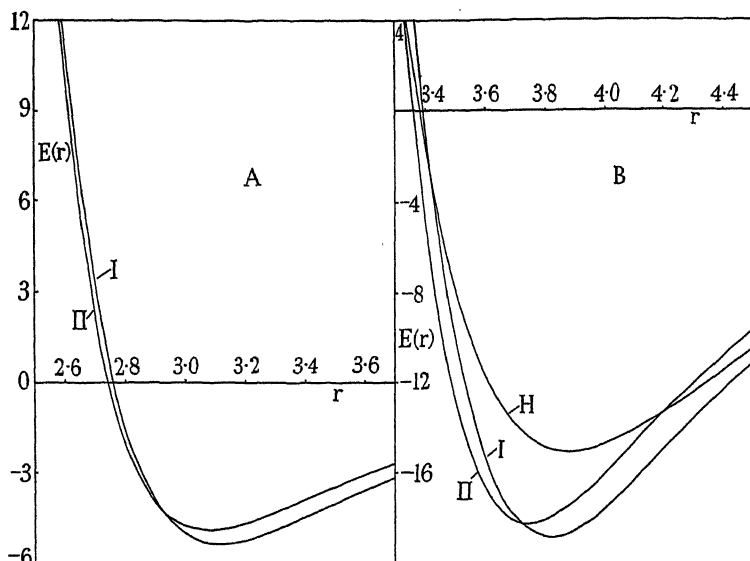


FIG. 2. Interaction of neon and argon atoms.

- A. Neon: I. $E(r) = be^{-r/\rho} - \mu r^{-6}$ (equation (26)).
 II. $E(r) = \lambda r^{-12} - \mu r^{-6}$ (constants from Table II).
 B. Argon: I. $E(r) = be^{-r/\rho} - \mu r^{-6}$ (equation (27)).
 II. $E(r) = \lambda r^{-12.5} - \mu r^{-6}$ (constants interpolated from Table II).
 H. Herzfeld interaction (equation (28)).

$E(r)$ is measured in units of 10^{-15} erg and r in Å.

giving $a_0 = 3.76$ Å and $U_0 = 2000$ cal./mole. The form of this is shown in fig. 2B, and for comparison $E(r) = \lambda r^{-12.5} - \mu r^{-6}$ (with constants interpolated from Table II), which leads to a similar value of U_0 . The relation between the two curves is the same as that previously noticed for neon and helium. The figure also shows an interaction suggested by Herzfeld and Goeppert-Mayer (1934), of the form

$$E(r) = 1.34 \times 10^{-9} e^{-r/0.345} - 111 \times 10^{-12} r^{-6} \text{ ergs}, \quad (28)$$

and based upon the properties of crystalline argon. Using equation (25) one finds from it $a_0 = 3.78 \text{ \AA}$ and $U_0 = 1670 \text{ cal./mole}$, a lattice energy which is much less than that assumed here, as is obvious from the shallower minimum to $E(r)$. The constant $\rho = 0.345$ is taken from Born and Mayer's work on alkali halide crystals, but it is unlikely to be correct for neutral atoms. Yet as pointed out by Hirschfelder, Ewell and Roebuck (1938) *this* function, as well as that of Lennard-Jones for $s = 14$ (which is very similar to curve II in fig. 2B), is in excellent agreement with the observed Joule-Thomson coefficients for argon, a fact which emphasizes the uncertainty attached to these interaction functions when gas data alone are investigated. The ambiguity is removed only when the properties of the crystal are taken into account: hence the importance of knowing these as accurately as possible.

From the use which has been made of a simple exponential to represent the repulsion potential it must not be concluded that this is the only or even the best function for the purpose. For example, over the comparatively short range which covers the points from which these analytic functions are derived a function $b_1 r^7 e^{-r/\rho_1}$ as suggested by Wasastjerna is equally as satisfactory as the simpler $be^{-r/\rho}$, and so presumably is any combination of exponential and polynomial. Wasastjerna (1932) has correlated the properties of the rare gases and of crystals containing ions of similar structure, using for the repulsive potential of two atoms or ions

$$b(\xi) = \frac{B}{\sigma_1 + \sigma_2} \xi^7 e^{-p\xi}, \quad (\sqrt{3} < \xi < 2.7) \quad (29)$$

in which $\xi = r/(\sigma_1 + \sigma_2)$, σ_1 and σ_2 are structural constants for the atoms or ions deduced from their refractivities, and B , p are universal constants to which Wasastjerna gives the values $5 \times 10^{-15} \text{ (erg cm.)}$ and 10 respectively.*

We have considered the possibility of using such a function for neon and argon in place of the simple exponentials given above. The necessary parameters have been evaluated for different values of μ , and that value of μ selected which leads to the best agreement with the crystal data. In

* In a later paper, Wasastjerna (1935) has pointed out that (29) is only a first approximation which neglects certain individual properties of the atoms. A more comprehensive expression is

$$b(\xi) = \frac{B}{\sigma_1 + \sigma_2} (a + \xi^7) e^{-p\xi},$$

in which a is characteristic of the pair of interacting atoms or ions.

this way the following interaction energies have been obtained, valid for the same ranges of r as (26) and (27):

$$\begin{aligned} \text{Neon: } E(r) &= 2.40 \times 10^{-9} r^7 e^{-r/0.147} - 9.0 \times 10^{-12} r^{-6}, & (30) \\ \text{giving } a_0 &= 3.07 \text{ A, } U_0 = 600 \text{ cal.} \end{aligned}$$

$$\begin{aligned} \text{Argon: } E(r) &= 2.11 \times 10^{-9} r^7 e^{-r/0.180} - 105 \times 10^{-12} r^{-6}, & (31) \\ \text{giving } a_0 &= 3.79 \text{ A, } U_0 = 1990 \text{ cal.} \end{aligned}$$

(r in Angstroms). Unfortunately these functions, almost identical in form with (26) and (27), cannot be combined in a general expression such as (29), which appears to be too crude to represent with sufficient accuracy the interactions of all the rare gases. The repulsive potentials given by (29) using Wasastjerna's values for σ (0.56 for Ne and 0.80 for A) are far from satisfactory. Some improvement is possible by suitable increases in the σ values, but even if these are adjusted to give the same value of p (~ 10) for both Ne and A, the corresponding values of B are then inconsistent. The argon function gives the better agreement; assuming $p = 10$, one obtains from (29), $\sigma = 0.90$, $B = 2.35 \times 10^{-15}$.

In these calculations, the possible effect of quantum corrections to the classical formula for the second virial coefficient has been neglected. For argon the correction is almost certainly small, but for neon it may be appreciable at the lowest temperature (65° K) and might account in part for the discrepancy mentioned above. However, this point must be reserved for a later investigation.

The writer would like to express his thanks to Professor R. H. Fowler for suggesting these calculations, to Professor J. E. Lennard-Jones for many discussions concerning them, and to the Department of Scientific and Industrial Research for a grant.

SUMMARY

The problem of deriving the form of the interaction of rare gas atoms from their observed second virial coefficients is attacked by the method introduced by Lennard-Jones, using the classical equation of state and for the interaction energy

$$E(r) = \lambda r^{-s} - \mu r^{-t}. \quad (1)$$

The method of calculation used by Lennard-Jones for helium, neon and argon, assuming λ and μ constant, has been modified and then extended to allow for a possible slow variation of λ with r . The bearing of the results

upon the true interaction of helium atoms is discussed, mainly in relation to the accuracy of the Slater field. For neon and argon it is emphasized that an unambiguous determination of the interaction energy requires an accurate knowledge of the properties of the crystalline form of these elements. The description of the repulsive potential by analytic expressions, such as $be^{-r/\rho}$ and $b_1r^7e^{-r/\rho_1}$, is considered, and parameters deduced which give the best correlation of gas and crystal data, on the assumption that (1) and kindred forms are valid for the interaction of neighbouring atoms only in the crystals.

REFERENCES

- Baber, T. D., and Hassé, H. R. 1937 *Proc. Camb. Phil. Soc.* **33**, 253.
 Bleick, W. E. and Mayer, J. E. 1934 *J. Chem. Phys.* **2**, 252.
 Born, F. 1922 *Ann. Phys., Lpz.*, **69**, 473.
 Born, M. and Mayer, J. E. 1932 *Z. Phys.* **75**, 1.
 Buckingham, R. A. 1937*a* *Proc. Roy. Soc. A*, **160**, 94.
 — 1937*b* *Proc. Roy. Soc. A*, **160**, 113.
 Clusius, K. 1929 *Z. phys. Chem. B*, **4**, 1.
 Fowler, R. H. 1936 "Statistical Mechanics", 2nd ed., ch. x.
 Gibby, C. W., Tanner, C. C. and Masson, I. 1929 *Proc. Roy. Soc. A*, **122**, 283.
 Herzfeld, K. H. and Goepfert-Mayer 1934 *Phys. Rev.* **46**, 995.
 Hirschfelder, J. O., Ewell, R. B. and Roebuck, J. R. 1938 *J. Chem. Phys.* **6**, 205.
 Holborn, L. and Otto, J. 1924–6 *Z. Phys.* **23**, 77; **30**, 320; **33**, 1; and **38**, 359.
 Huggins, M. L. 1937 *J. Chem. Phys.* **5**, 143.
 Keesom, W. H. and Taconis, K. W. 1938 *Physica*, **5**, 161.
 Kirkwood, J. G., and Keyes, F. G. 1931 *Phys. Rev.* **37**, 831.
 Lennard-Jones, J. E. 1924 *Proc. Roy. Soc. A*, **106**, 463.
 — 1931 *Proc. Phys. Soc.* **43**, 461.
 Lennard-Jones, J. E. and Ingham, A. E. 1925 *Proc. Roy. Soc. A*, **107**, 636.
 Nijhoff, G. P., IJin, B. and Keesom, W. H. 1928 *Commun. Phys. Lab. Univ. Leiden*, 188c.
 Ruhemann, B. and Simon, F. 1932 *Z. phys. Chem. B*, **15**, 389.
 Simon, F. and v. Simson, C. 1924 *Z. Phys.* **25**, 160.
 Slater, J. C. 1928 *Phys. Rev.* **32**, 349.
 de Smedt, J. and Keesom, W. H. 1926 *Commun. Phys. Lab. Univ. Leiden*, 178b.
 de Smedt, J., Keesom, W. H. and Mooy, W. H. 1930 *Commun. Phys. Lab. Univ. Leiden*, 203 E.
 Uhlenbeck, G. and Beth, E. 1936 *Physica*, **3**, 729.
 Wasastjerna, J. A. 1932 *Soc. Sci. Fenn. Comm. Phys.-Math.* **6**, Nos. 18–22.
 — 1935 *Soc. Sci. Fenn. Comm. Phys.-Math.* **8**, No. 20.

The formation of negative ions by positive-ion impact on surfaces

BY R. H. SLOANE, D.Sc. AND R. PRESS, M.Sc.

Queen's University, Belfast

(Communicated by E. V. Appleton, F.R.S.—Received 12 July 1938)

INTRODUCTION

The quantal theory of negative ions has now been developed considerably (Massey and Smith 1936; Massey 1938), but on account of difficulties of computing it is usually necessary to assume rather than to prove that a given ion exists, and then to discuss the probability of its formation by different processes. The work described here is a contribution to the experimental side of this subject. It had its origin in a projected attempt to measure the capture cross-section of mercury atoms for electrons, as a verification of Massey and Smith's (1936) then unpublished theory of this process. In considering this it became clear that the experiment would be one of unusual difficulty. Before proceeding with it, therefore, it was decided to verify the existence of Hg^- as a stable entity, which is assumed in the quantal theory. This was in doubt since Stille (1933), in some careful experiments, had recently failed to obtain it from the plasma of various forms of discharge through mercury vapour, although it is known that negative ions tend to accumulate in such regions (Emeleus and Sayers 1938). Whilst one of us was repeating his experiments, with modifications which led to essentially the same results and will be described elsewhere, Arnot and Milligan (1936*b*) reported that they had obtained Hg^- by bombardment of metal surfaces with Hg^+ . A comparison of their work with our own showed only one essential point of difference, namely that the construction of their apparatus did not permit of degassing *in situ*, a condition satisfied with our tubes. Both for this reason and because of the intrinsic importance of their discovery it was thought desirable to repeat part of their work, with apparatus geometrically similar in electrode construction, but capable of being degassed in a furnace under vacuum. We were again unable to obtain Hg^- after the apparatus had been degassed and in operation for a short time, although initially a complex negative-ion beam with mass of the order of that of Hg^- was obtained. There were, however, always present several light negative ions, which had the excess energies found by Arnot

and Milligan (1936*b*) with Hg^- . The conditions under which these were formed led us to suppose that they were produced by bombardment of the metal surfaces by mercury positive ions (Press 1937; Sloane 1937) and not by capture of electrons by positive ions of the same species, the process suggested by Arnot and Milligan (1936*b*). The existence of a process of this type, which may be conveniently termed "sputtering" (Sloane and Press 1938), is also implied by some earlier work by J. S. Thompson (1931) which has, so far as we know, never been published in detail. It is, however, impossible to decide definitely that a particular positive ion (e.g. Hg^+) is liberating another kind of negative ion (e.g. CO^-) when it hits the surface, so long as one is producing the negative ions on a metal surface in a plasma or ionization chamber. One cannot overlook the possibility of the negative ion being formed from its own positive ion, since the latter may be present in the plasma or ionization chamber, and CO^+ , for example, was in fact shown in our work to be there from the positive-ion mass spectra, although in quantity only a fraction of 1 % of Hg^+ . An unambiguous decision on this point can only be reached by first isolating a particular positive ion by a mass spectrograph, then bombarding a surface by this in a *high vacuum*, and finally making a mass spectrographic and energy distribution analysis of the resultant negative ions. We have built a double mass spectrograph for this purpose and find that negative ions of one kind possessing energies in excess of that imparted to them by the accelerating fields can be produced by bombardment with positive ions of another kind (Sloane and Press 1938).

An account of the experiments with the single mass spectrograph is given in §1. The experiments with the double mass spectrograph are described in §2.

1. PRODUCTION OF NEGATIVE IONS BY POSITIVE-ION IMPACT

(a) *Experimental arrangements*

The construction of the first tube used is shown in fig. 1. It consisted of an ion-generating system (FS_2), a magnetic deflexion chamber (T) and a detecting cylinder (C). It was blown in pyrex, with tungsten seals for the electrode leads to permit of degassing. The ions were produced in the region G_2A and accelerated to S_2 . All the electrodes here were made of nickel, except the two filaments F (one of which is shown) and their supports, which were of tungsten and nickel-chrome respectively. The nickel and nickel-chrome had been subjected to preliminary degassing in a

vacuum furnace. The anode (A) was a cylinder closed at one end by a disk and covered at the other by a gauze G_1 . A cylinder was welded to G_1 and reached nearly to the gauze G_2 , which was welded to a second cylinder open on the left-hand side. The electrode S_1 was a short cylinder closed by a disk with a slit 2 by 8 mm. in the centre, and was pressed out of a single piece of nickel. On to this disk a nickel cylinder of smaller diameter was welded

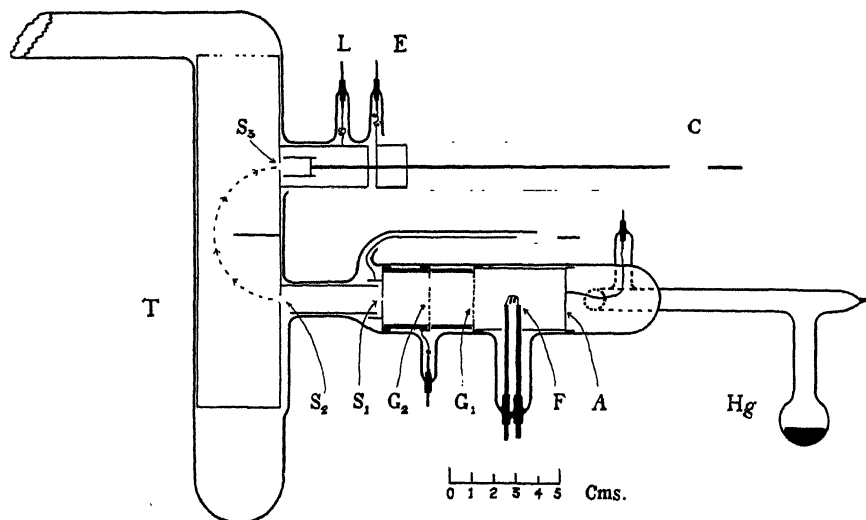


FIG. 1. First tube used for generating negative ions.

which served to cover the right-hand side of a nickel-chrome electrostatic screening channel connected to S_2 . The electrodes S_1 , G_2 and G_1 were spaced by pyrex cylinders. The filaments were 0.2 mm. in diameter with a potential drop of less than 2 V across each. The cylinder containing S_1 pressed closely against the glass walls. The tube from S_2 to A was shielded magnetically by a close-fitting external jacket of soft iron, 0.7 cm. thick. The mercury vapour pressure to the right of S_1 was controlled by the temperature of the mercury reservoir.

The analysing chamber T was supported between the tapered pole pieces (7 cm. diameter) of an electromagnet, and was lined with a non-magnetic nickel-chrome cylinder, 4 cm. in diameter, with two slits (S_2 , S_3) 5.95 cm. apart, of dimensions 2.5 by 8 mm. and 2 by 8 mm. respectively. To maintain a low vapour pressure in T and in the space between S_1 and S_2 , the lower end of T was immersed in liquid oxygen, while the upper end was connected through a liquid oxygen trap to a three-stage mercury-vapour

pump. The whole tube was degassed under vacuum in a resistance furnace at 450° C for many hours.

The electrical arrangements were such that a voltage V_0 could be applied between F and G_1 , a voltage V_1 between G_1 and G_2 , and a voltage V_2 between G_2 and S_1 . The energy of the ions entering T was controlled by the single voltage V_3 , applied between S_2 , and G_1 , G_2 or F , depending on whether positive or negative ions were being analysed. With these arrangements, whatever the interelectrode fields, the total energy given to the ions emerging from S_2 was measured by V_3 . In operating the mass spectrograph, variation of V_3 was usually preferred to a magnetic field variation. The ion currents to the Faraday cylinder connected to C were measured on a Compton electrometer which could be operated with ease at a sensitivity of 10,000 mm./V, and was provided with auxiliary capacities and parallel resistances.

The deflecting system was calibrated by beams of positive ions from a mercury-vapour discharge passing between F and A , and the calibration checked from the geometry of the apparatus and the fields. With a magnetic field of a little over 3000 gauss, the maximum permissible with this design of apparatus because of the wide spacing of the pole pieces and the consequent large stray magnetic field, the Hg^+ ion was brought into the Faraday cylinder with an accelerating voltage (V_3) of 20 V. We have consistently found Hg_2^+ to be present, in agreement with Arnot and Milligan (1936a) as well as Hg^+ and Hg^{++} . The calibration was also of interest in showing the presence, in spite of the precautions taken, of some light ions. The most prominent, assuming they were singly charged, had masses of about 28 and 39. Although an observed complexity in the peak due to the former may indicate the presence of second-group hydrocarbons (C_2H_x^+), the metallurgy of nickel makes it probable that this peak was largely due to CO^+ . Results obtained by Aston (1933) for metal carbonyls make it almost certain that the heavier ions of mass about 39 are mainly C_3H_3^+ . No light ion had an intensity greater than 0.3 % of that of Hg^+ , so that the discharge in FA must have been predominantly through mercury vapour.

(b) Types of negative ions

To generate negative ions either the gauze G_2 , or the filament F and its supports, were bombarded with positive ions from the plasma in the ionization chamber FA . Fig. 2 shows a typical negative-ion mass spectrum obtained by bombarding G_2 with the various positive ions present, using the arrangement of potentials shown on the figure. The temperature of the mercury was 18° C and the current from F 11.4 mA. There is no sign of a

peak in that range of values of V_3 (20 V) which corresponds to an ion of mass comparable to that of Hg^- , which was explored repeatedly; a current of 10^{-16} amp. could have been measured. Negative ions of this mass were only obtained from F and then only in small quantity, in the earlier runs. The beam was complex. It was not studied in detail because the maximum current never exceeded 2×10^{-14} amp. and had decreased to less than 10^{-16} amp. permanently after the tube had been in use for a day. The peak current fell to half value in an hour with a discharge of 25 mA from F passing continuously, although no observable change took place in the discharge or its associated circuits.*

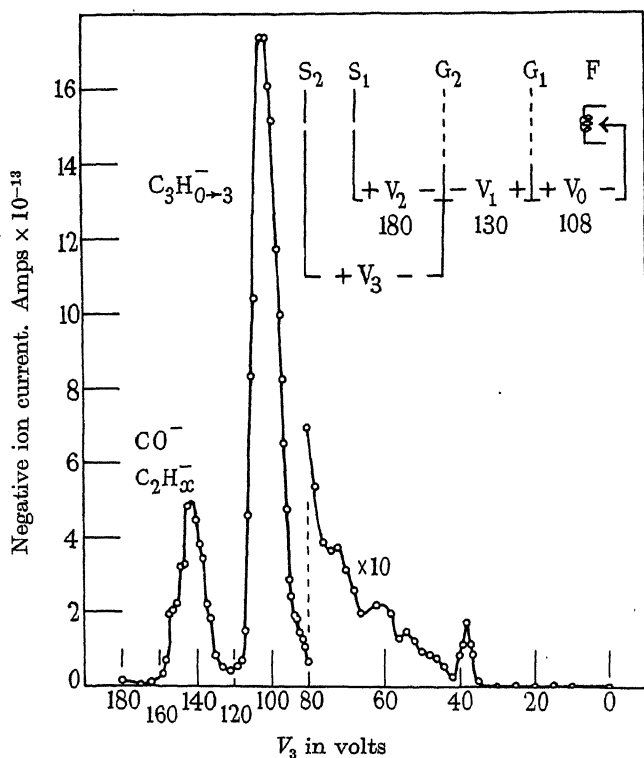


FIG. 2. Negative ions from G_2 .

It is clear, however, from fig. 2 that light negative ions were present in our tube. Some of these have been identified as indicated, but the identifications are not quite certain in detail because of the poor resolving power

* Under apparently comparable conditions of purity Stille (1933) found that the heavy ions reappeared on contaminating the mercury with hydrogen; this did not occur with our tube after hydrogen had been admitted through a palladium regulator.

of this form of mass spectrograph. For convenience the main ones will be referred to as CO^- and C_3H_3^- . The C_3H_3^- peak is higher than the CO^- peak, whilst in the positive-ion spectrum the CO^+ peak was much higher than the C_3H_3^+ peak. The essential features of both positive- and negative-ion spectra were verified repeatedly, and it was immaterial whether they were obtained by variation of V_3 or the magnetic field. The diffuse peaks in the region of V_3 from 50 to 80 V may be due to Cr^- and Ni^- , a tentative identification rendered plausible by the heavy disintegration which the filament and its nickel-chrome supports were found to have undergone when the tube was opened. The mass number of the unidentified sharp peak to the right of these is about 104.

(c) *Energy distribution and place of origin*

Arnot and Milligan (1936*b*) found that a negative ion produced on a surface by bombardment by the corresponding positive ion might have an energy in excess of that imparted to it by the accelerating electric fields. It was found immediately that the light negative ions, which we suspected were being formed by bombardment of the surfaces with Hg^+ , had the same property. It is important for establishing any theory of this excess energy to be sure that it is not produced through any macroscopic property of the discharge, in particular through oscillations. The need for caution in this respect is emphasized by the well-known production of electrons with excess energy in this way (Penning 1926; Tonks and Langmuir 1929). We have given some attention to this point.

For studying the excess energies we have used what is essentially the retarding potential method previously used by Arnot and Milligan (1936*b*). Particles of a given kind differing in momentum, due to their leaving their place of origin with a range of initial velocities, can be brought on to the slit S_3 by variation of either the magnetic field (H) or the accelerating voltage (V_3). Previous workers have varied H , but we have preferred to vary the accelerating voltage, thereby eliminating effects due to change in any stray magnetic field in FA . At the same time no electrical disturbance of sheaths in FA occurs, since V_3 is varied by alteration of the potential difference between S_1 and S_2 (fig. 2). The total number of ions approaching S_3 which leave the source (F or G_2) with more than any particular initial energy V_r can be found by making the Faraday cylinder above S_3 , V_r volts negative to the source, and varying V_3 so that the whole, imperfectly focussed beam moves across S_3 laterally. Arnot and Milligan found that the peak heights of the curves so obtained (in their case by variation of H) were proportional to the areas beneath the peaks, and we have assumed that the

same is true in our case. The distribution curve for the initial energy of the ions is obtained by first plotting a curve showing peak height against V_r , and then differentiating this with respect to V_r . Examples of the various curves, which will be referred to again later, are shown in figs. 3, 4 and 7. Fig. 7*a* shows a typical series of peaks obtained for various values of V_r , and demonstrates immediately the existence of excess energies up to at least 30 V. Fig. 3 is an example of the first curve obtained from curves of the type of fig. 7*a*, and figs. 4 and 7*b* are examples of the final energy distribution curves. The energy distribution curves for positive ions can be obtained in the same way by appropriate reversal of fields. They tail off in the opposite sense because of a slight amount of scattering between FA and S_3 . The amount of scattering in our tubes, as judged by the tail of the Hg^+ curve, was about the same as in Arnot and Milligan's case, although we have used much smaller accelerating voltages. The negative-ion distribution curves have also low energy tails, but the course of the curves on this side was erratic in some cases. Since both the CO^- peak in the mass spectrum (fig. 2) and, when plotted on a larger scale, the $C_3H_3^-$ peak show a complex structure, it is probable that overlapping of the low energy side of the main peak and the high energy tail of one due to a less massive ion was responsible for this irregularity. The particular retarding potential and energy distribution curves shown in figs. 3 and 4 respectively were obtained for $C_3H_3^-$ ions generated on G_2 , but all the energy distribution curves obtained were similar wherever the ions were formed.

The fact that ions are generated with excess energies at both G_2 and F , where the discharge conditions are entirely different, makes it improbable that plasma oscillations are responsible for the production of the excess energies. It leaves open, however, the possibility of another type of oscillation coming into play. When ions are being generated at G_2 under conditions similar to those which gave the curves in figs. 3 and 4, all the positive ions passing through G_2 are turned back by V_2 and, if they do not strike the side of G_2 facing the slit, are caused to describe orbits about the wires of G_2 , with which they ultimately collide. When the tungsten filament is acting as source, a similar orbital motion can take place around it. The orbital motion around the filament is not readily controlled, but that around the gauze G_2 , which is similar to that of electrons in a Barkhausen-Kurz type of oscillator,* can be controlled by the voltage V_2 . Fig. 5 shows a plot of the positive-ion current to G_2 against V_2 when V_1 was held constant at 130 V. It shows that for V_2 less than 80 V the total positive-ion current to G_2 was approximately

* Very recently the possibility of devising an ionic oscillator of this type has been investigated by Kownacki and Ratcliffe (1938).

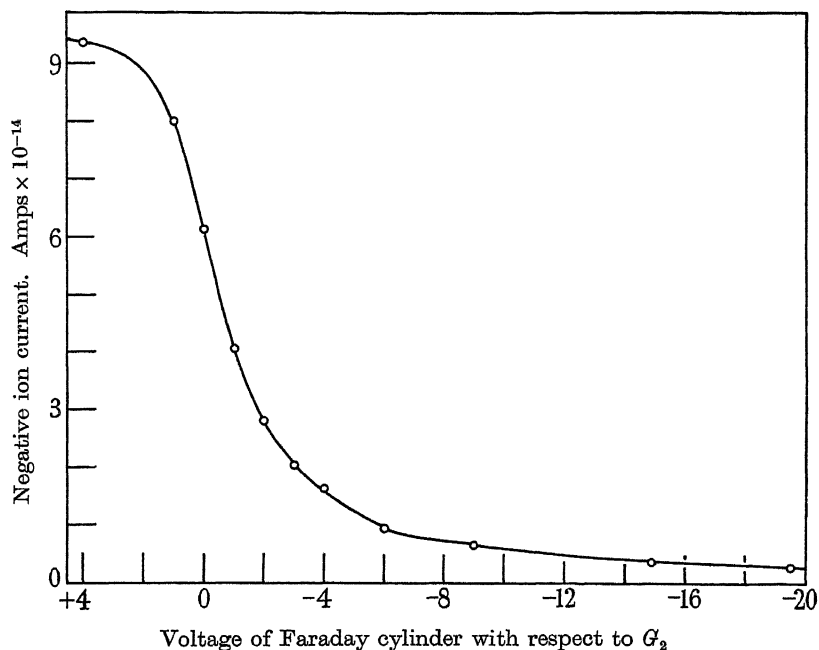


FIG. 3. Retarding potential curve for $C_3H_3^-$ ions generated on G_2 . $V_0=85$ volts, $V_1=130$ volts, $V_2=180$ volts, and $V_3=103$ volts when V_r was zero. Current from $F=6.5$ mA. Temperature of Hg = 18° C.

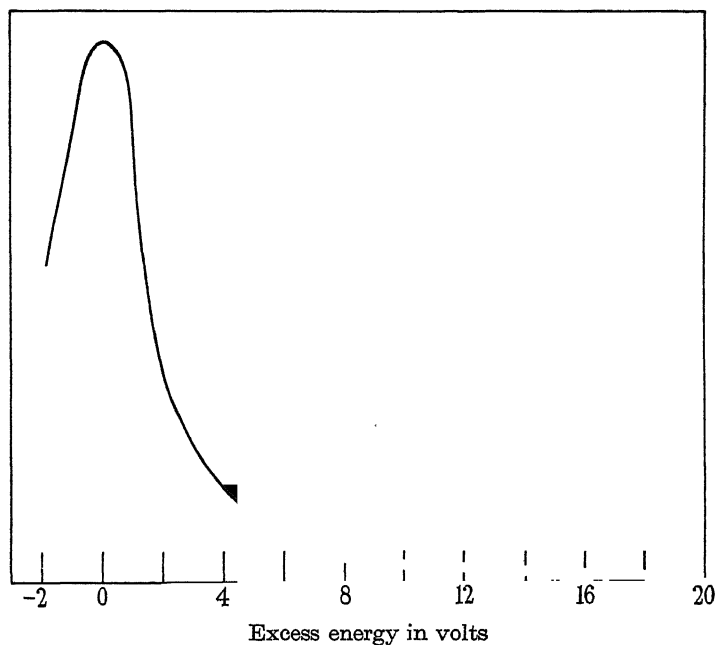


FIG. 4. Energy distribution curve derived from fig. 3.

constant, with a slight tendency to increase as V_2 increased, which may be attributable to some positive ions reaching the edges of the cylinder of which the gauze was an end covering. For V_2 greater than 80 V the positive ions passing through G_2 were turned back to an extent depending on V_2 . With V_2 greater than 130 V all the positive ions passing through G_2 were turned back by V_2 and some were made to move in orbits about the wires. To examine whether the possibility of these oscillations was of significance, retarding potential curves were obtained under the conditions marked by α , β , and γ in fig. 5. The negative-ion currents with $V_r = 0.0$ V were in the ratio 4 : 5 : 3. The energy distribution curves proved to be all of the same shape as that shown in fig. 4, which was obtained under conditions closely similar to γ . We conclude that Barkhausen-Kurz oscillations of the ions are not responsible for the excess energies.

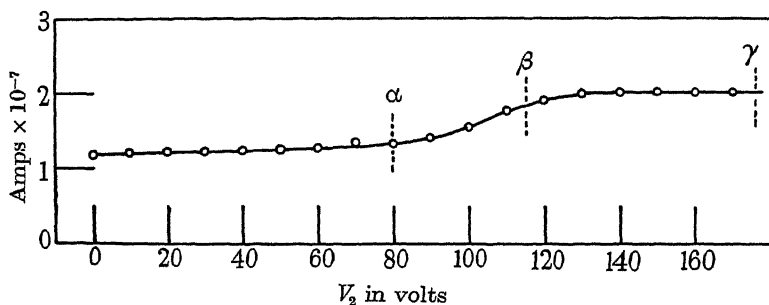


FIG. 5. Positive ion current to G_2 . $V_0 = 85$ volts, $V_1 = 130$ volts, current from $F = 4.75$ mA.

The formation of negative ions with excess energies on the gauze G_2 , under the conditions α , when the positive ions were mainly passing straight through the gauze in the direction of S_1 , raises the question as to whether the negative ions are not produced exclusively by glancing blows of the positive ions on the wires. To examine whether negative ions can be formed by normal impact of positive ions and at the same time to confirm the absence of any effect of Barkhausen-Kurz oscillations the modified generating system shown in fig. 6 was built. The wire filaments of the first chamber were replaced by two oxide-coated cathodes C . These were welded to the nickel cylinder, in which they were enclosed, by short thin leads which simultaneously provided good electrical contact and satisfactory thermal insulation. The spacing of the cathodes from the cylinder was such that under the conditions of operation the positive-ion sheath round the walls of the cylinder was continuous with those on the cathodes. For the generation of negative ions the electrode which contained S_1 was used as anode for the

discharge from C , and the disk D was bombarded with positive ions by application of a potential difference (V_0) between it and S_1 . The energy given to the negative ions generated on D was that acquired from the potential difference V_3 between S_2 and D . When D was held negative to the plasma the impact of the positive ions on it was normal except for a small current at the edges,* which were not opposite the slits, and there is no possibility of circulation of the ions round D . Fig. 7 shows the peak heights and energy distribution curve obtained in a retarding potential run for $C_3H_3^-$ ions drawn off from D under the conditions stated. These results which are again almost identical with those obtained with the gauze and filament in the first tube seem to eliminate finally the possibility of any macroscopic electrical effects.

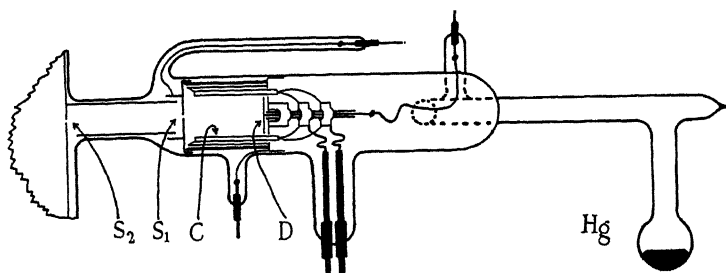


FIG. 6. Ionization chamber for generating negative ions on a flat disc.

We do not feel justified from our results in making any estimate of the efficiency of production of negative ions in terms of energy of the positive ions. In the case of generation at the gauze G_2 , since appreciable negative-ion currents are recorded under conditions α (fig. 5), it does not appear to us justifiable to assume, as has been done by previous workers, that the effective positive-ion current under conditions γ is half of the total current to the effective part of G_2 . In the case of generation at D again we do not know what fraction of the recorded positive-ion current generates the negative ions removed through the slits.

2. DOUBLE MASS SPECTROGRAPH

(a) *Significance of previous results*

Arnot and Milligan (1936*b*) suggested that a negative ion is formed by a positive ion capturing two electrons from a surface, and rebounding with part of its initial energy, in some cases with molecular dissociation (Arnot

* The shape of the positive-ion sheath on a negatively charged disk of finite size has been described by Langmuir and Mott-Smith (1924).

1937*a*). R. A. Smith (1938) has succeeded recently in fitting a quantal theory to this process. We felt that this was unlikely to be the explanation of the origin of the light negative ions in our apparatus, particularly since

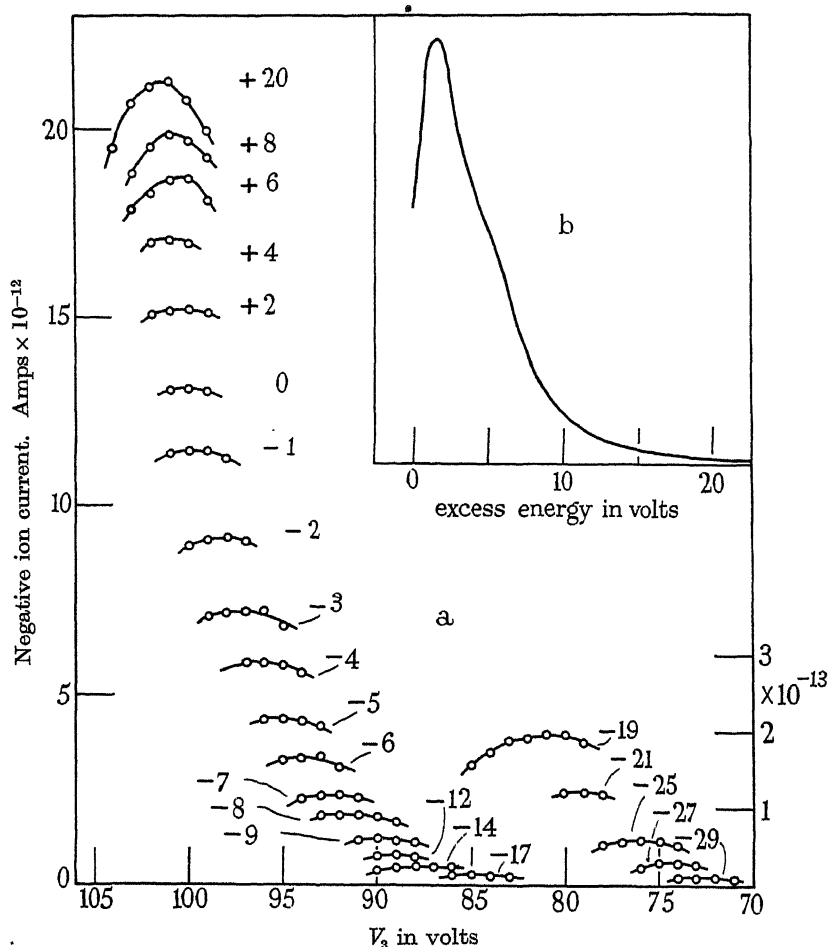


FIG. 7. (a) Peak currents for $C_3H_3^-$ ions generated on the disk *D* (fig. 6). The number against each peak gives the potential in volts of the Faraday cylinder with respect to *D*. $V_0 = 175$ volts, arc potential $V_a = 26$ volts, arc current = 160 mA, positive ion current to *D* = 8.67×10^{-4} amp., and temperature of Hg = 34°C .

(b) Corresponding energy distribution curve, uncorrected for contact potential difference.

the mass spectrum showed that Hg^+ was present in far greater concentration than any other ion. A more probable process seemed to us to be that bombardment of the surface by the Hg^+ (and as later work showed Hg^{++}) ions was liberating material occluded there in the form of electronegative

films (Sloane and Press 1938). This view received some support from Arnot and Milligan's curve showing the negative-ion yield as a function of the energy of incidence of positive ions, which is generally similar to the curves showing the rate of sputtering of neutral cathode material in terms of the energy of the positive ions (Langmuir and Kingdon 1923). This does not preclude a given positive ion from producing its own negative ion, by sputtering, nor indeed the existence of Arnot and Smith's process as an alternative mode of origin of negative ions.

With regard to the existence of Hg^- , Arnot and Milligan's experiments certainly provide strong evidence for its reality, but we feel that this point cannot be decided definitely without an investigation of the negative mass spectrum with an instrument sufficiently precise for an accurate determination of mass and particularly for the resolution of the isotopes.

In absence of a more detailed knowledge than it seems possible to obtain about the constitution of the plasma in *FA* (fig. 1) and the relative yields of different negative ions, it appears to be impossible to decide, beyond all doubt, with any such generating apparatus, that one positive ion incident on the surface can give rise to a negative ion of different species.* This can only be done with certainty by the double mass spectrograph.

(b) Experimental arrangements

The apparatus is necessarily more complicated than that previously described, but it could be designed with confidence for normal bombardment of a disk in view of the behaviour of the tube of fig. 6. Its essential features are shown in fig. 8.

Since we were primarily interested in the process of formation and not in the types of ions, and were moreover analysing both beams, it was considered unnecessary to construct it so that it could be degassed. The rectangular deflexion box, 9 mm. wide, was made of a chromium-plated brass frame with massive iron faces which served as pole pieces for the electromagnet. The surfaces between the brass and iron were ground accurately, held together with set screws, and made vacuum tight with Apiezon grease. Five iron ground joints were screwed into the brass frame as shown. The two lower were connected to the pumps and a liquid oxygen trap respectively. Of the three upper the centre one supported the positive-ion generating system, the left-hand one the nickel-chrome disk to be bombarded, and the right-hand one a Faraday cylinder for detection of negative ions from the disk. Slits, baffles, internal and external guard rings

* See, however, Arnot and Beckett (1938).

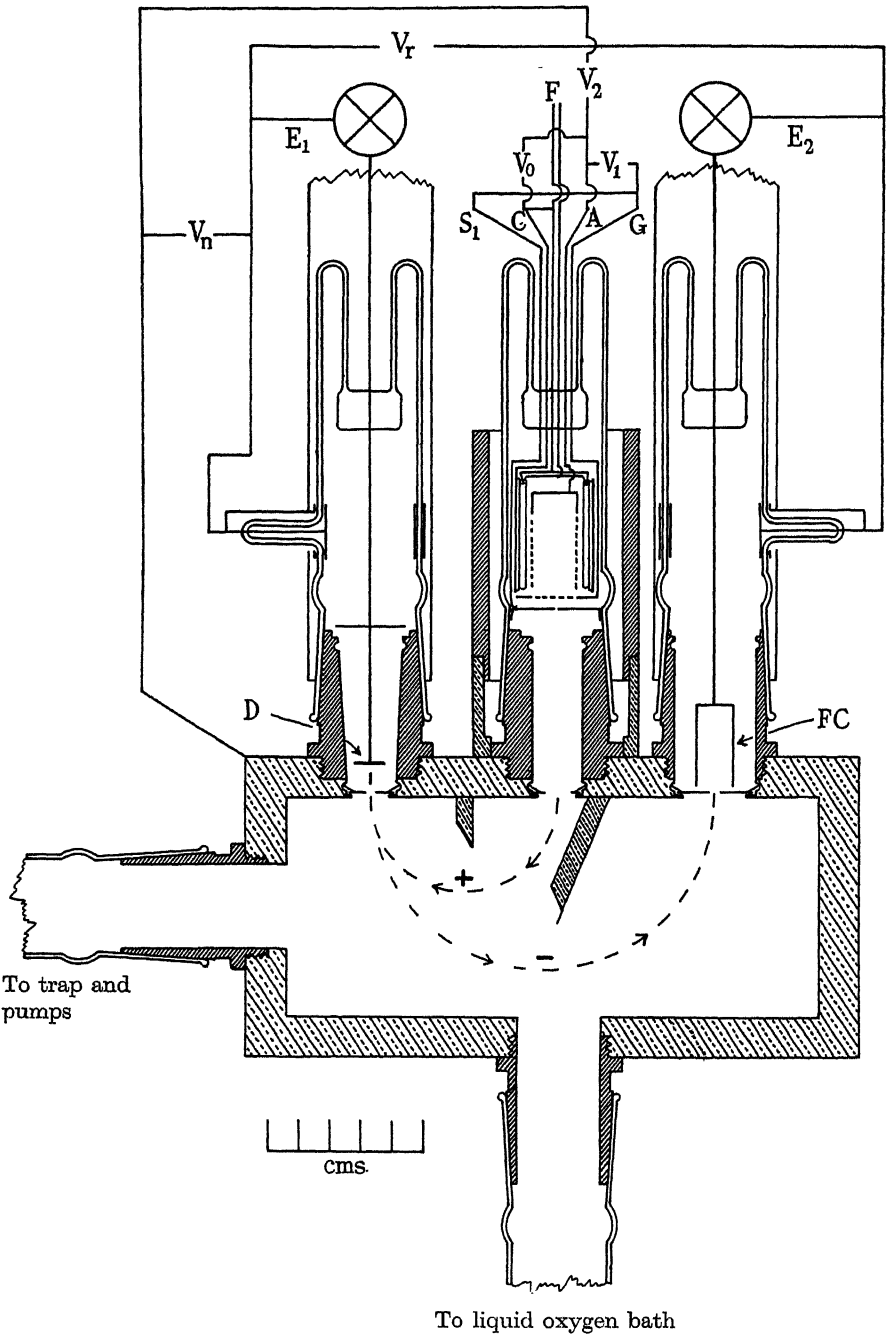


FIG. 8. Double Mass Spectrograph. Iron shaded with full lines, brass with alternate broken lines.

and electrostatic screens were provided as shown. A mercury reservoir which is not shown was attached to the centre tube. The iron ground joints also served as magnetic shields and were so arranged that they did not complete the magnetic circuit of the electromagnet. There was a further external magnetic shield for the middle system. Water could be circulated between the lower part of this and the middle ground joint to keep the latter cool. The positive-ion generating system consisted of four indirectly heated oxide-coated cathodes (two of which are shown) with heaters connected in parallel to F and emitting surfaces attached to a cylinder, as in the tube of fig. 6, and connected to C . A discharge was passed, at potential difference V_0 , between the cathodes and the central gauze cylinder connected to A . Positive ions, generated inside the last-mentioned cylinder, were withdrawn by the application of a negative potential V_1 between A and a flat gauze connected to G , and were further accelerated through the slits beneath this. The positive ions were deflected on to the disk D in the left-hand tube and the current to D was measured by the electrometer triode E_1 , with a flexible parallel resistance circuit. The energy of the positive ions in the deflexion chamber was controlled by V_2 and was usually of the order of 100 V. They were further accelerated by a potential difference V_n of up to 4000 V between the slit of the left-hand tube and D , so that the latter was bombarded by positive ions of energy $V_b = V_n + V_2$. The potential difference V_n , together with any initial energy possessed, determined the energy with which negative ions from the disk passed downwards into the deflexion chamber. Their subsequent path to the right-hand tube where they were collected in the Faraday cylinder is indicated. The current to the cylinder was measured by the Compton electrometer E_2 used before. With this apparatus it is essential to have the deflexion box earthed since it is connected to the electromagnet. The electrometers and associated circuits must therefore be all at a high potential with respect to earth and operated by remote controls.

The energies possessed by the positive and negative ions are not independent since they must traverse the same magnetic field. The procedure adopted, in view of this fundamental limitation, was as follows: Setting the magnetic field to a convenient value between 6000 and 10,000 gauss, V_2 was adjusted so as to bring any desired positive ion to the slit of the left-hand tube. To discover if this could generate a given negative ion on the disk the accelerating potential V_n within the left-hand tube had to be set at such a value that the negative ion would be swung round in the fixed magnetic field to the Faraday cylinder in the right-hand tube. To obtain a negative-ion spectrum, V_n had therefore to be varied, which meant that the bom-

barding energy of the positive ions could not be held constant without elaborate readjustment of fields.

On dismantling the tube the focus spot on D could be seen as a discoloured spot in the shape of a sharply defined elongated ellipse.

(c) *Results*

This type of apparatus opens up a wide field for investigation. So far we have confined our attention mainly to showing with it that negative ions of one kind can be produced by bombardment of the disk D with positive ions of another kind. The positive ions used have been Hg^+ , Hg^{++} and a complex beam with mass of approximately 28, presumably mainly CO^+ . Each ion has been found to produce a spectrum of negative ions. Both Hg^+ and Hg^{++} produced several ions with masses between 18 (lighter have not yet been studied) and 120. Using Hg^+ a careful search was made for Hg^- with a negative result. The most prominent negative ion produced by Hg^+ and Hg^{++} had mass 28 and was presumably CO^- . This and other ions were also produced by CO^+ . The negative-ion spectra produced by Hg^+ and Hg^{++} were very similar; the spectrum due to CO^+ has not yet been mapped in detail.

If the negative ions are being formed on D by bombardment with positive ions as supposed, then the current of negative ions of a given species into the Faraday cylinder should be proportional to the positive-ion current to the disk, at least when the latter is small. For larger positive-ion currents the electro-negative film might be removed by bombardment at the focus spot at a greater rate than it could be renewed by diffusion from other parts of the surface or the body of the metal, and the efficiency of production might fall. Some typical results for formation of CO^- are shown in fig. 9. The positive-ion current was varied in each case by variation of conditions in the ion-generating tube. It was measured with V_n applied, and therefore includes any secondary electron emission which may occur from D ; this may however be expected to be proportional to the positive-ion current so long as the energy of the positive ions is constant. The bombarding energy of the Hg^+ ions was 2630 V, that of Hg^{++} 2560 V and that of CO^+ 3068 V. In each of these cases there is proportionality between the two currents.

Since the bombarding voltages were not greatly different these curves give a rough idea of the relative efficiency of production of CO^- by the different positive ions. Ignoring differences in secondary electron emission from D , which may be different for different ions, the yields per incident particle are approximately in the ratios 1 : 0.5 : 0.1 for Hg^+ , Hg^{++} and CO^+ . It must be remembered however that each Hg^{++} ion has approxi-

mately twice the energy of an Hg^+ ion. The construction of the apparatus does not allow of any reliable estimate of the absolute probability of formation of a negative ion by a given positive ion.

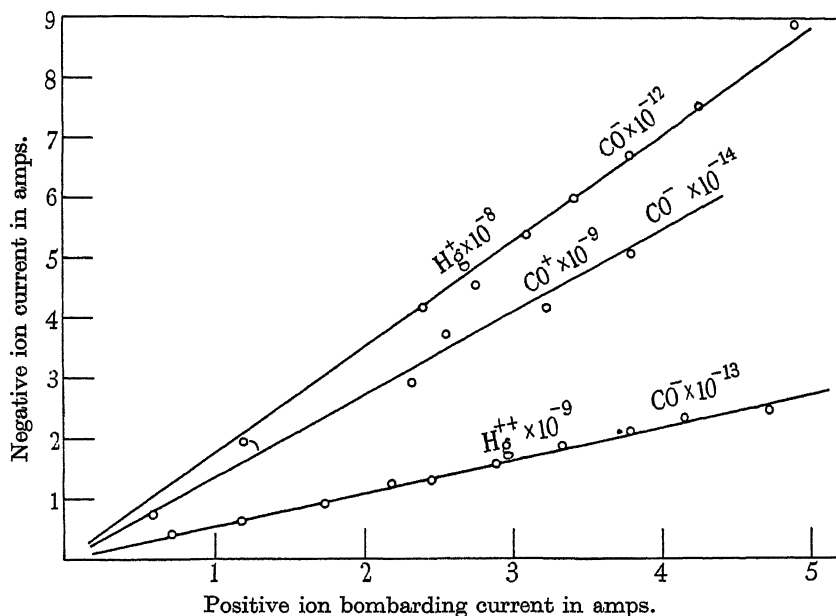


FIG. 9. Proportionality between bombarding positive ion currents and resultant negative ion currents.

(d) Retarding potential measurements

Further verification that these negative ions are produced on the disk is provided by showing that they have excess energies. To obtain the energy distribution of the negative ions as they are emitted from the disk, a retarding potential V_r (fig. 8) was applied between the disk and the Faraday cylinder, as with the earlier tubes. The requirements for obtaining good energy-distribution curves with this apparatus are exacting. Since the negative-ion currents recorded near the tail of the peak are only of the order of 10^{-14} amp. and less, the Compton electrometer must be worked near its maximum sensitivity. This increases the time required to get each reading, with the result that a considerable time is required to get a full retarding potential run. During this time all fixed fields and currents must be maintained accurately constant.

The energy distribution of the CO^- ions produced by bombardment of D by a current of 4.33×10^{-8} amp. of Hg^+ at 2570 V was investigated by this

method. It was found that the current of negative ions to the Faraday cylinder was 1.6×10^{-12} amp. with zero retarding potential, 2.4×10^{-13} amp. with 14 V retarding potential and 6.0×10^{-14} amp. with 18 V retarding potential. These measurements show that some negative ions leave *D* with initial energy and that their energy distribution has approximately the same form as the curve of fig. 7.

The CO^- ions produced by bombarding with Hg^{++} have not yet been investigated in this way. It has however been noticed that the position of the peak is slightly displaced towards lower values of V_n in the mass spectrum when Hg^{++} is substituted for Hg^+ , conditions being otherwise the same except for the small difference in bombarding voltage consequent on the change in V_2 necessary to bring Hg^{++} into the slit below *D*. This indicates that the majority of the CO^- ions produced under these conditions by Hg^{++} have appreciable excess energy, provided that the displacement is not due to a change in the type of negative ion.*

These results seem to leave no room for doubt that bombardment of a metal surface by positive ions of one kind can liberate from it negative ions of another kind.

The writers wish to express their thanks to Professor K. G. Emeleus and Dr H. S. W. Massey for their interest in this work; to Mr J. Wylie, B.A., and Mr Buick for constructing the Compton electrometer and electromagnet, and the metal work of the double mass spectrograph to our specification; to Dr J. D. Cockcroft for the loan of the patterns for the electromagnet; and to Dr C. C. Paterson, Director of the Research Laboratories of the General Electric Company for the oxide-coated cathodes. One of us (R. Press) also acknowledges gratefully the award of a maintenance grant from the Government of Northern Ireland.

SUMMARY

It has been found that when a hot-cathode discharge is passed through mercury vapour, light negative ions are produced from surfaces exposed to positive-ion bombardment. The negative ions include some with energies greater than could be imparted to them in the existing electric fields. Experiments are described which eliminate the possibility that the excess energies could be produced through electrical oscillations. The conditions under which the negative ions appear indicate that they are produced

* Or, just possibly, to interaction between the two beams near S_3 .

mainly through sputtering of occluded electronegative films under the action of mercury positive-ion bombardment, but data from experiments of this type cannot be interpreted unambiguously. The existence of this process has been shown directly by a double mass spectrograph in which beams of positive ions of known kind are allowed to impinge normally on a metal disk in high vacuum and are found to produce a spectrum of negative ions, with properties similar to those of the negative ions obtained in the previous experiments. The existence of Hg^- as a stable entity seems to be still in doubt.

REFERENCES

- Arnot 1937*a* *Proc. Roy. Soc. A*, **158**, 137.
 — 1937*b* *Proc. Roy. Soc. A*, **158**, 157.
 Arnot and Beckett 1938 *Nature, Lond.*, **141**, 1011.
 Arnot and Milligan 1936*a* *Proc. Roy. Soc. A*, **153**, 359.
 — — 1936*b* *Proc. Roy. Soc. A*, **156**, 538.
 Aston 1933 "Mass-spectra and Isotopes." London: Arnold.
 Emeleus and Sayers 1938 *Proc. Roy. Irish Acad.* **44**, 87.
 Kownacki and Ratcliffe 1938 *Nature, Lond.*, **141**, 1009.
 Langmuir and Kingdon 1923 *Phys. Rev.* **22**, 148.
 Langmuir and Mott-Smith 1924 *Gen. Elect. Rev.* **27**, 616.
 Massey 1938 "Negative Ions." Camb. Univ. Press.
 Massey and Smith 1936 *Proc. Roy. Soc. A*, **155**, 472.
 Penning 1926 *Physica*, **6**, 241.
 Press 1937 Thesis, Belfast.
 Sloane 1937 Thesis, Belfast.
 Sloane and Press 1938 *Nature, Lond.*, **141**, 872.
 Smith 1938 *Proc. Roy. Soc. A*, **168**, 19.
 Stille 1933 *Ann. Phys., Lpz.*, **17**, 635.
 Thompson 1931 *Phys. Rev.* **38**, 1389.
 Tonks and Langmuir 1929 *Phys. Rev.* **43**, 195.
-

mainly through sputtering of occluded electronegative films under the action of mercury positive-ion bombardment, but data from experiments of this type cannot be interpreted unambiguously. The existence of this process has been shown directly by a double mass spectrograph in which beams of positive ions of known kind are allowed to impinge normally on a metal disk in high vacuum and are found to produce a spectrum of negative ions, with properties similar to those of the negative ions obtained in the previous experiments. The existence of Hg^- as a stable entity seems to be still in doubt.

REFERENCES

- Arnot 1937*a* *Proc. Roy. Soc. A*, **158**, 137.
 — 1937*b* *Proc. Roy. Soc. A*, **158**, 157.
 Arnot and Beckett 1938 *Nature, Lond.*, **141**, 1011.
 Arnot and Milligan 1936*a* *Proc. Roy. Soc. A*, **153**, 359.
 — — 1936*b* *Proc. Roy. Soc. A*, **156**, 538.
 Aston 1933 "Mass-spectra and Isotopes." London: Arnold.
 Emeleus and Sayers 1938 *Proc. Roy. Irish Acad.* **44**, 87.
 Kownacki and Ratcliffe 1938 *Nature, Lond.*, **141**, 1009.
 Langmuir and Kingdon 1923 *Phys. Rev.* **22**, 148.
 Langmuir and Mott-Smith 1924 *Gen. Elect. Rev.* **27**, 616.
 Massey 1938 "Negative Ions." Camb. Univ. Press.
 Massey and Smith 1936 *Proc. Roy. Soc. A*, **155**, 472.
 Penning 1926 *Physica*, **6**, 241.
 Press 1937 Thesis, Belfast.
 Sloane 1937 Thesis, Belfast.
 Sloane and Press 1938 *Nature, Lond.*, **141**, 872.
 Smith 1938 *Proc. Roy. Soc. A*, **168**, 19.
 Stille 1933 *Ann. Phys., Lpz.*, **17**, 635.
 Thompson 1931 *Phys. Rev.* **38**, 1389.
 Tonks and Langmuir 1929 *Phys. Rev.* **43**, 195.
-

A Discussion
on
Plastic flow in metals

12 May 1938

Opening address

BY PROFESSOR W. L. BRAGG, F.R.S.

A perfect metal crystal is extremely weak, yielding to a vanishingly small shearing force. As deformation proceeds the perfection of the crystal structure is destroyed, and the resistance to shear is increased. After severe cold-working in all directions the metal attains a steady state, with random orientation of the crystal grains, and characteristic mechanical properties.

Yield takes place along definite crystalline planes, and parallel to definite directions in these planes (e.g. [111] and [110] in crystals with a face-centred cubic lattice). A rod under tension develops slip-bands parallel to these planes. Deformation is not uniform, but is concentrated in bands of macroscopic dimensions separated by undeformed portions, more slip-bands being formed as extension proceeds.

These characteristics of plastic flow, revealed by the work of Andrade, Polanyi, Schmid, Taylor, Orowan and others, must be explained by theory; the theories hitherto put forward are admittedly incomplete and diverge widely in their nature. The contrast between a brittle material and one which undergoes plastic flow can be explained in a qualitative way. The cohesion of a brittle material depends upon the juxtaposition of certain atoms in the structure arranged in a definite way, and if this arrangement is destroyed the scheme of the crystal is broken down. In a metal, on the other hand, the metal ions are held together by the common free electrons. Any scheme of arrangement in which the metal ions are approximately equidistant and closely packed satisfies the main conditions for cohesion; the many alternative "phase structures" in alloys is evidence of this. Hence, when the metal is distorted the atoms can pass through a continuous series of transitional configurations without breakdown of the whole structure. Some such process must be taking place on a slip-band, and it leaves the structure less perfect than before. What we should like to have is more precise information about the atomic movements when slip is

taking place. The conception of atomic planes slipping over each other is probably far too simple. All theories agree in supposing that some kind of dislocation appears at one point in the crystal and travels across it. What is the nature of this dislocation? Is it centred around a point, or along a line, or, as is more probable in the author's view, is it a dislocation over a surface? When the metal is severely cold-worked, and has reached a state with characteristic mechanical properties, what determines this final state? Various lines of evidence point to a mosaic of crystallites of dimensions $1-5\mu$ inside each crystal grain. How do these crystallites confer hardness on the material? Can we develop a quantitative theory to explain the observed values of the mechanical characteristics? These are the problems before us.

Dr C. H. DESCH, F.R.S. When a mass of metal is plastically deformed, the deformation is frequently greater in some regions than in others, producing such effects as Lüders lines and the "déformations banales" of Osmond. The direction of these deformed zones is determined by the direction of stress only, and is independent of the crystal structure. In an aggregate of crystal grains, similar effects are seen within the grains. A specimen of cupro-nickel (an α -solid solution with face-centred cubic lattice), after being reduced 80 % by rolling, shows when etched a number of dark lines, in some grains parallel with the direction of rolling, in others inclined at about 40° (Adcock 1922). These bands are regions of greater strain, as shown by their greater chemical activity and by the fact that new crystal grains originate solely in them on annealing. Dark bands also appear in β -brass on compression, and these have sometimes been found to disappear in the course of years by healing of the deformed region. They appear to be distinct from the very straight bands produced in the same alloy by deformation, which have been identified (Greninger and Mooradian 1938) as formed by transformation to the α -phase by strain. In both instances, however, a localization of strain is indicated. Similar effects are produced during fatigue. A polished surface shows an increasing number of slip-bands as fatigue proceeds, and in the later stages accumulations of such bands are seen at wider intervals. True slip-bands disappear on polishing and etching, but such accumulations leave a permanent effect, shown by the more rapid attack of the reagent. These markings were originally described as cracks (Ewing and Humfrey 1902), but they are revealed by etching before the stage of cracking is reached. The conclusion is drawn that the plastic flow of a grain in a metallic aggregate (observations on

single crystals do not seem to be available) is localized, certain bands, the direction of which depends on that of the applied stress, being much more highly deformed than the neighbouring parts of the crystal. The spacing of such bands is wide in comparison with that of slip-bands, and therefore very large compared with the lattice dimensions. It is suggested that this fact should be taken into account in any theory of plastic flow which involves the breakdown of the crystal lattice into small fragments. The breakdown in the bands must be either greater, or the lattice deformation greater, to account for the increased chemical activity.

Professor G. I. TAYLOR, F.R.S. The geometry of the mechanism contemplated in W. L. Bragg's model, i.e. the movement of a boundary between two crystals at slightly differing orientations, is identical with that of a plane of dislocations, all of the same sign, moving across a crystal. The stress in the material on either side of the plane is not that which would be calculated by adding together the stresses due to each of the dislocations as if it existed alone. This is because in Bragg's model the whole material on one side of the plane is rotated relative to that on the other. In a single crystal a plane of dislocations, as in Bragg's model, could move across a crystal producing the same resulting plastic strain as that which would be due to the passage of randomly distributed dislocation, which was the arrangement contemplated in my model. The elastic strain energy associated with Bragg's model is very much less than in my model, and this is a feature which is much in favour of the former, because the energy calculated as necessarily associated with my model is greater than is observed when the internal energy associated with plastic strain is released on heating. On the other hand it is difficult to imagine how the displacements of the material which are necessarily associated with Bragg's model could accommodate themselves so that boundaries of crystal grains would remain in contact during plastic straining.

Professor N. F. MOTT, F.R.S. *Resistance to glide in alloys.*

The object of this note is not to present any new results about the theory of plastic flow in metals. It is rather to suggest a line of experimental and of theoretical work which, it is hoped, might lead to more precise information about the forces resisting plastic flow, and hence about hardness.

We start from a consideration of Taylor's theory of strain hardening (Taylor 1934*a*). In this theory the following assumptions are made:

(1) That slip takes place by the motion through the crystal of "dislocations" (fig. 1). For our purpose it does not matter whether the dislocations exist already in the crystal, or whether they are formed under the action of the strain at external or internal surfaces. The important assumption is that in a crystal free from internal strain, a dislocation will start to move under a very small external strain.

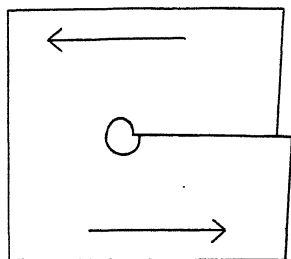


FIG. 1

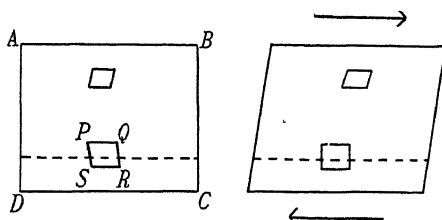


FIG. 2

(2) That the cause of hardness in crystals is the existence of internal strain. In fig. 2 (a) $ABCD$ represents a solid possessing internal strains, so that a small block of the material $PQRS$, which would take the form of a cube if cut out of the block, has the distorted form of a parallelogram shown in the figure. Clearly, if a stress is applied as shown by the arrows, no dislocation can travel along the dotted line until the strain is sufficient to bring $PQRS$ back to the form of a cube (fig. 2 (b)).

We shall adopt this hypothesis here, and shall not discuss the rival hypothesis that hardness is due to surfaces of misfit between small crystals which, according to this hypothesis, are supposed to obstruct the passage of a dislocation.

(3) In the strain hardened crystals the hardness is due to the strain caused by the dislocations themselves, which travel only a certain distance and then become stuck. As the strain increases, so does the number of stuck dislocations. The strain due to these impedes the motion of further dislocations.

Of the assumptions made by G. I. Taylor, the third has probably the least general validity, since internal strain could be produced in many other ways during the process of slip, e.g. by local melting and recrystallization. The first two assumptions, on the other hand, may well be valid not only for strain hardened, but for age hardened materials, and for alloys

in general. It is the purpose of this note to examine the possibilities of extending the theory in this direction.

It is believed that in an age hardened alloy consisting, for instance, of two metals A and B , atoms of B initially in solid solution in A form aggregates in the crystal lattice of the original alloy $A + B$. These aggregates will tend to contract or expand to the atomic volume characteristic of the metal B , and so set up a state of strain in the surrounding medium. If one knew the size of a given aggregate, and if one made some assumption about its shape, it would obviously be possible to calculate this state of strain.

If a metal were a two-dimensional aggregate of atoms one could then deduce the hardness. One would have to consider all the "lines" along which the metal could slip; along any one of these lines the strain would vary in an irregular way, and the passage of a dislocation could not take place until the external strain were so great that the internal strain were everywhere in the same direction. But in three dimensions the state of affairs is more complicated. A Taylor dislocation (fig. 1) has the form of a line in the plane of slip perpendicular to the direction of slip. In fig. 1 the dislocation must be thought of as extending right through the crystal in a direction perpendicular to the plane of the paper. Now if we consider the *average* internal strain in an age hardened alloy along any such line, it will certainly be zero; at some points of the line the strain will be in one direction, and will tend to help the passage of the dislocation; at other points it will be in the opposite direction, and will tend to hinder it. Thus if we use the Taylor theory in its present rather two dimensional form, we cannot correlate internal strain and resistance to shear.

It seems to us, however, unlikely that dislocations, of atomic dimensions in the direction of slip, will be formed extending for a distance of the order of millimetres in the perpendicular direction. It seems more probable that slip starts at crystal surfaces by the motion of dislocations of finite length only. The type of deformation envisaged* is shown in fig. 3, which is drawn in the slip plane. The shaded area is supposed to have slipped a distance equal to one atomic diameter. We assume that, once slip of this type has started, the length l of the dislocation increases, and eventually perhaps the heat generated allows a slip of macroscopic dimensions of the observed type to begin.

If now we knew the length l that a dislocation must have if macroscopic slip is to start, then we could correlate the hardness of a material with its internal strain. The only way to find this out seems to the author to be by means of theoretical and experimental work on age hardened alloys. In the

* I am indebted to Dr Orowan for discussions on this point.

neighbourhood of one of the aggregates referred to above, the internal strain will be large, and will fall off with the distance. The maximum *mean* strain that can exist along a line l will thus depend markedly on the length l . By relating the observed external strain necessary to produce slip to the calculated mean internal strain, it ought to be possible to obtain an estimate of l .

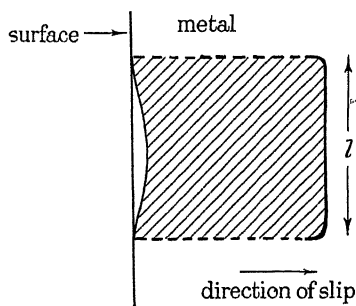


FIG. 3

Dr E. OROWAN. *The rate of plastic flow as a function of temperature.*

Information about the mechanism of a chemical reaction can be derived from the temperature dependence of its velocity. If α is the energy barrier that must be overcome in an elementary reaction, the velocity of the macroscopic reaction is given by

$$Ce^{-\alpha/kT}, \quad (1)$$

C being independent of temperature or very much less dependent than the exponential factor.

From this formula the energy barrier or its molal value, the activation energy $A = N\alpha$, can be obtained if the velocity as a function of temperature is known.

There is a close analogy between chemical reactions and plastic deformation. In any kind of solid, amorphous or crystalline, plastic deformation proceeds in distinct steps in which only a relatively small number of atoms are involved, and these elementary processes of deformation are activated by thermal agitation. Thus from the temperature dependence of the rate of plastic flow the activation energy can be obtained; the dependence of the activation energy upon the stress gives, on the other hand, valuable information concerning the mechanism of the plastic flow. This may be demonstrated by some examples.

(1) *Amorphous plasticity.* This type of plastic flow is not restricted to

amorphous bodies but includes also, for example, the intergranular plasticity of polycrystalline metals. Its mechanism consists of changes of place within groups of two or more molecules or atoms. Such atomic rearrangements give rise to elastic strains in the body that cause minute changes of its external shape. In the plastic range of temperature rearrangements take place incessantly, without respect to whether the body is under external stress or not. In the unstressed state the changes of the external shape produced by different rearrangements will cancel. If, however, the body is stressed, then the probability of a rearrangement increases if its contribution to the macroscopic strain agrees with the external stress, and decreases in the opposite case. Consequently, plastic flow will set in.

Let n be the number of all possible rearrangements; α_i the energy barrier of a particular rearrangement denoted by the number i ; $\delta_i = b_i \sigma$ the perturbation of α_i under the influence of the external stress σ (b_i is a constant, since in practical cases σ is of a lower order of magnitude than the cohesive stresses); ϵ_i the change of a certain strain component due to the rearrangement i ; and ds/dt the time rate of the strain component in question. Application of (1) leads then to the formula

$$\frac{ds}{dt} = \sigma \frac{1}{kT} \sum_{i=1}^{i=n} \epsilon_i b_i C_i e^{-\alpha_i/kT} = f(T) \sigma, \quad (2)$$

representing Newton's law of viscosity for amorphous solids.

If the activation energies α_i are densely grouped around a mean value α , this formula can be approximated by

$$\frac{ds}{dt} = \sigma \frac{\text{const.}}{T} e^{-\alpha/kT}. \quad (3)$$

Thus the activation energy of amorphous flow is practically independent of stress.

(2) *Crystalline plasticity.* In crystals the atoms interact very strongly with each other and stable plastic displacements, however small, must affect a great number of atoms. Correspondingly, in the unstressed state the energy barrier is very high, and plastic deformation cannot set in unless the barrier is lowered by applying a shear stress. Thus the activation energy is here a steep function of the external stress. An approximative expression representing the activation energy of plastic gliding, given by R. Becker (1925; cf. Orowan 1934) is

$$\alpha = \frac{(\sigma_0 - \sigma)^2 V}{2G}; \quad (4)$$

here σ_0 , V and G are constants and σ is the shear stress in the glide plane resolved parallel to the glide direction.

(3) *Creep*. The scientific study of creep began with the work of Professor Andrade (1910, 1914) who found that the rate of creep contains two components, one dying away with time ("primary creep"), the other being independent of time ("secondary creep"). He attributed the primary creep to the deformation of the crystal grains, the secondary (or "viscous") creep to an amorphous kind of deformation.

Now recent researches have shown that it is the time rate of gliding but not the glide strain itself that is determined by the stress. In other words, plastic gliding in crystals is, as well as the flow in amorphous bodies, of an essentially viscous nature; that this had not been recognized long ago, was due to the fact that, in the case of gliding, the "viscosity" increases in course of the deformation, sometimes very rapidly, by strain hardening.

Consequently, the viscous component of creep cannot be ascribed immediately to amorphous flow and further evidence is necessary for ascertaining its mechanism.

A study of the rate of secondary creep as a function of stress cannot decide this question, since there is neither proportionality to the stress, nor any similarity to the plastic behaviour of the material in short-time tests.

The problem can be solved, however, by calculating the activation energy from the temperature dependence of the secondary creep rate. With regard to the complexity of the creep rate-stress curves it is surprising that, in a considerable number of cases, the secondary creep rate as a function of temperature obeys a formula (1) with the activation energy independent of stress (Kanter 1937). This proves that in these cases, in agreement with the assumption made by Professor Andrade, secondary creep is a manifestation of an amorphous type of plasticity.

Taking into account this result and some other observations relating to creep, we arrive at the following picture of the mechanism of creep. After applying the load gliding starts at places with a high local stress (primary creep); it becomes, however, increasingly checked by the mutual hindrance of the crystal grains, by the interference of different sets of active glide planes, and, to a smaller extent, by the deterioration of the lattice in the glide zone. Finally plastic deformation can only proceed according to the rate at which these obstacles are removed by atomic rearrangements activated by thermal agitation (secondary creep).

Since the mutual hindrance of neighbouring crystal grains, the interference of different sets of glide planes, and the lattice deterioration inside the glide zones together represent the strain hardening, the constancy of

the activation energy means that secondary creep is a flow by strain hardening recovery (thermal softening).

Professor E. N. da C. ANDRADE. One of the striking features of the plastic flow of metals, to which Professor W. L. Bragg has alluded, is that, for a given lattice, the glide takes place on planes parallel to a certain crystallographic plane and in a certain crystallographic direction. Another fundamental phenomenon is the hardening that accompanies glide. I wish to bring forward certain new experimental results bearing on these points and to indicate certain regularities which may help towards the formation of a satisfactory theory.

There is wide agreement that the glide direction is always the most closely packed direction, but the glide plane is not always the most closely packed plane. With body-centred crystals different planes have been found to act as glide planes with different metals. With regard to α -iron much uncertainty prevails: (110), (112) and (123) have all been given as glide planes by different workers. With tungsten (112) is the glide plane: with β -brass, under certain conditions (G. I. Taylor 1928), (110) is operative.

In conjunction with Mr L. C. Tsien and Miss Y. S. Chow (L. C. Tsien and Y. S. Chow, 1937, for molybdenum: other result unpublished) I have undertaken work to find the glide planes for body-centred metals at various temperatures. With molybdenum the glide plane is (110) at 1000° C and (112) at 300 and 20° C: with sodium it is (123) at 20° C, (110) at -80° C and (112) at -185° C: for potassium it is (123) at 20° C. These results, and those of previous observers on body-centred metals, can be generalized if we express the temperature as $\theta = T/T_m$, where T is the temperature at which the stress is applied and T_m is the melting point on the absolute scale. We then have that for high values of θ the glide plane is (123): for medium values (110): for low values (112). In all cases the glide plane as determined experimentally lay very near the plane with which it has been identified: intermediate positions were not found. Fig. 4 shows our actual determinations on a stereographic projection, and the results, our own as well as those of other workers on β -brass and tungsten, are embodied in the following table. Iron has not been included on account of the uncertainty that still prevails for this metal.

Metals	θ	Glide plane
W, Mo, Na	0.08-0.24	(112)
β -brass, Mo, Na	0.26-0.50	(110)
Na	0.80	(123)

It must be remarked, however, that, owing to the high symmetry of the body-centred system, it has seldom happened that the resolved shear stress on the actual glide plane was markedly smaller than that on the other two planes which can act as glide planes at other temperatures. In several cases the resolved shear stress on two of the planes in question, sometimes on all three, was about equal. There seems no doubt as to the preference expressed in the table, but we cannot say, from our results, that the preference was very marked. It appears, rather, that all three planes are well adapted for glide, and that temperature exercises a small determining influence.

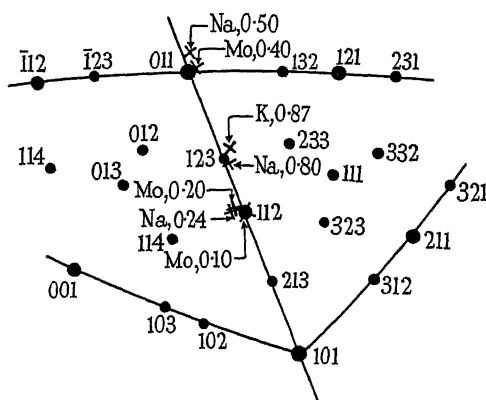


FIG. 4

A closely connected problem is that of the regularity of the glide surfaces. G. I. Taylor and C. F. Elam (1926) have shown that with α -iron the nature of the traces of the slip-bands, whether straight or irregular, can be explained by supposing that the metal slips not in slices, but in rods, of which the section may be, but need not be, hexagonal. This would lead to straight line traces on a plane parallel to the direction of slip, but not on any other plane. Their experiments were carried out with blocks, subject to compression. In stretched wires this would mean that the traces would appear straight when the bands are viewed along a line normal to the wire axis and in the plane of slip, but irregular from any other point of view. In particular the heads of the glide traces would be very irregular. With very pure iron ("Ommet" iron, for which a purity in excess of 99.98% is claimed—see Espe and Knoll 1936), Y. S. Chow and I have obtained slip-bands showing little irregularity at the head (fig. 5). This would seem to indicate that impurities may exercise a determining influence as to regularity or irregularity of surfaces of slip. In this connexion attention may be directed to the experiments of Greenland (1937), who showed that with the purest mer-

curry, regular or irregular slip surfaces could be produced at will. Any slight preliminary bending of the crystal wire led to perfectly plane surfaces, whereas absolutely unstrained wires gave irregular surfaces.

The irregular surfaces may be explained by glide taking place simultaneously on several planes through the glide direction, the actual glide surface being made up of small planes, bounded by lines parallel to the glide direction, succeeding one another at irregular distances in such a way as to form one surface. The influence of impurities and of strain are complex and require further consideration, but once more the existence of three possible different families of glide planes, the operation of which is determined by small subsidiary influences, may be taken as a basis.



FIG. 5. Slip-bands with very pure iron. (a) Head of bands; (b) bands seen from the side.

Another great problem, to the solution of which Professor Taylor's contributions are well known, concerns the hardening which takes place when single crystals are strained. It seems to be generally true that strain hardening is accompanied by an effect which manifests itself as the transformation of the circular spots of a Laue photograph into the elongated figures which are known as asterisms. These are, for instance, very well marked in iron, for which an extension of 16 % at atmospheric temperature produces the effect shown in fig. 6. The asterisms have, in certain cases, been shown to be due to a rotation of crystal fragments about a line in the glide plane perpendicular to the direction of glide (e.g. W. G. Burgers 1931; Andrade and Tsien 1937). A single crystal which has been strained is no longer a single crystal: in particular the treatment of a single crystal of high melting point, which has been considerably strained at atmospheric temperature, as a single crystal is likely to lead to erroneous results. We can reasonably suppose that the rotation of the crystal fragments takes place in the glide planes (better called, perhaps, glide lamellae, since they are regions very many atoms in thickness), which are regions of great strain,

and that the hardening may be attributed to this. Such a picture is based on the theory that the perfect lattice is very strong, and that glide takes place owing to flaws, from which the disturbance is propagated in the manner suggested by G. I. Taylor (1934*a, b*). If the crystallites are strong unflawed fragments they will, in their rotated positions, exercise a locking action.



FIG. 6. Pure iron extruded by 16 % at atmospheric temperature; asterisms.

In sodium the distorted spots do not appear as continuous smears, but rather as a line of discrete spots (Andrade and Tsien 1937). The angle of rotation corresponding to the extreme spots in such a line is easily measured, and may be taken as a measure of the amount of crystal break up that has taken place in the glide planes. Miss Chow and I have carried out a series of experiments in which sodium wires are extended by a fixed amount. The range of angle of rotation of the crystallites, $\Delta\theta$, is measured from the extreme spots, and the corresponding hardening is also observed. This is done at various temperatures. We find that at low temperatures, where the hardening corresponding to given strain is pronounced, $\Delta\theta$ is much greater than at higher temperatures, where the same strain produces much less hardening. For instance, 5.5 % extension produces, at temperatures of -185 , -80 and 20°C , values of $\Delta\theta$ of 9, 5 and 2.5° respectively. The experiments are being continued, but already seem to be sufficient to establish

that the rotation of the crystal fragments is not determined by the strain, but by the stress.

I suggest, then, that hardening is inevitably accompanied by crystal break up, and rotation of the fragments in the glide lamellae, and that in metals near the melting point, where there is little hardening, it will be found that there is little rotation, even at large strains.

G. D. PRESTON. One very characteristic feature of plastic strain is the production of work-hardening in the strained material. Without suggesting that age and strain-hardening are identical it is still possible that the two phenomena have some points of similarity. I would like to amplify Professor Mott's remarks on this matter by briefly describing the results of some experiments recently completed on an age-hardening alloy which indicate one type of structural change which produces an increase in hardness. The material investigated was an alloy of high purity aluminium with 4 % by weight of copper. When quenched from 500° C the copper is all in solution, being distributed at random on the points of the aluminium lattice; in this state the solid solution is metastable and the hardening takes place at room temperature and at temperatures up to about 250° C as a result of the efforts of the alloy to reject the copper from solution. If the heat treatment is prolonged at elevated temperatures the hardness reaches a maximum and then diminishes. For the purpose of the present discussion the point of interest is that X-ray photographs of single crystals, which have had a heat treatment sufficient to produce the maximum degree of hardness, exhibit a peculiarity, in the form of a set of lines associated with the normal reflexions, which can be interpreted as indicating that the copper atoms are becoming segregated on certain (100) planes in the crystal. The thickness of these regions is small, perhaps no more than one or two atomic distances, while their area is relatively large. The experiments do not tell us whether the surfaces to which the copper atoms migrate exist, as surfaces of "misfit" of the sort described by Professor Bragg, before the accumulation of copper on them, and it is not clear whether the hardening is due to the production of a new flaw or to the healing of an existing one. In the latter event the hardening would be due to the production of greater perfection in the crystal and an approach to its theoretical strength; it would then have to be admitted that age-hardening was a different process from strain-hardening, where destruction of perfection is taking place.

Finally, I would like to allude to the work of Dehlinger and Graf (1930) and of Schäfer (1933) on the effect of plastic strain on alloys in the ordered

state. Briefly it may be said the effect of cold work on an alloy such as Cu_3Au in the ordered state is to reduce it to a state of disorder. This might be interpreted as evidence that during strain high temperatures are locally produced, but whether such temperatures are in fact attained it is clear that cold work leaves an ordered alloy in the state that would be produced by exposure to a temperature of 600 or 700° C followed by quenching.

Dr W. H. HATFIELD, F.R.S. Whilst much interested in Dr Bragg's remarks, I am not clear upon his ideas regarding the difference between brittle and plastic materials. Some metals which are brittle have still a high degree of cohesion, and, judged from their electrical conductivity, must have free electrons, just as in the case of the plastic metals. A further point of interest is that there are many phase structures in alloys which are extremely brittle, although they have simple space lattices, e.g. 9–12 % nickel steels which are brittle and are mixtures of body- and face-centred cubic lattices. The properties of different steels of similar hardness, but dissimilar toughness also indicate that there is surely more in the subject than simple phase structure.

The stress to determine slip is in considerable measure governed by crystal size, and the presence, nature and location of the atoms of added elements. Under the effect of slip it is stated by Dr Bragg that the deformation is not uniform, but is concentrated in bands of macroscopic dimensions separated by undeformed portions. Whilst admitting that this is the prevalent view, I feel that the meticulous examination would disclose that the apparently undeformed portions are very seriously modified in properties, and are not undeformed portions, as suggested.

The cohesion of brittle materials seems to him to depend on the juxtaposition of certain atoms arranged in a definite way. He then states that if this arrangement is destroyed the scheme of the crystal is broken down. It is not clear what is meant by this. I consider that the breaking down of the crystal under plastic deformation results in debris which still remains crystalline in nature. If Dr Bragg suggests the creation of what is in effect the liquid phase which spontaneously recrystallizes, I would agree.

There are other aspects which require consideration. For instance, it is known that austenitic steels (face-centred cubic lattice) in general work harden much more rapidly than a steel in the ferritic condition (body-centred cubic lattice). Why should the disturbance in the vicinity of the slip plane, and hence the degree of cold work hardening, be different in the

two structures? Why also is there such a large difference in the rate of cold work hardening of different austenitic steels?

Incidentally, it is a matter of great interest that the effect of cold work in austenitic face-centred cubic lattice is to transform a proportion of the material into body-centred cubic lattice, and it surely is part of the general problem to include an explanation of the effect of cold work in achieving such a result. As an illustration I exhibit a photogram (fig. 7) prepared by Dr A. Westgren in March 1925 in an attempt by us to explain magnetic properties induced by cold work into austenitic material.

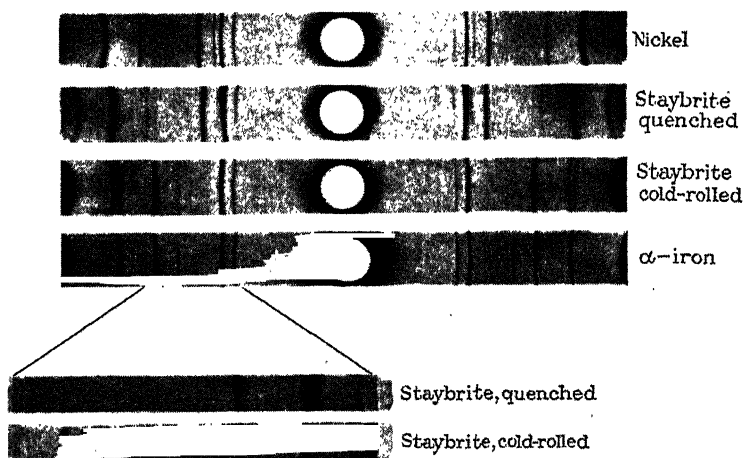


FIG. 7. Powder photographs of quenched and cold rolled staybrite (18 % Cr, 9 % Ni) compared with photographs of nickel and α -iron, chromium, K -radiation. Diameter of cylinder, 55 mm.

Dr Bragg makes the statement that when material is severely cold worked it reaches a stage with "characteristic mechanical properties", what are these characteristic mechanical properties? Would not the general subject be more readily advanced by a survey of the existing information relative to the effect of cold work on the physical constants of the material, i.e. the decrease of specific gravity, the increase of electrical resistance, the change in the coefficient of expansion.

In conclusion, there seems at the moment to be no effective theory available in explanation of the phenomenon of plastic flow and its influence on metals, and it is to be hoped that the discussion will be continued.

REFERENCES

- Adcock, F. 1922 *J. Inst. Met.* **27**, 73.
 Andrade, E. N. da C. 1910 *Proc. Roy. Soc. A*, **84**, 1.
 — 1914 *Proc. Roy. Soc. A*, **90**, 329.
 Andrade, E. N. da C. and Tsien, L. C. 1937 *Proc. Roy. Soc. A*, **163**, 1.
 Becker, R. 1925 *Phys. Z.* **26**, 919.
 Burgers, W. G. 1931 *Z. Phys.* **67**, 605.
 Dehlinger, U. and Graf, L. 1930 *Z. Phys.* **64**, 359.
 Espe, W. and Knoll, M. 1936 "Werkstoffe der Hochvacuumtechnik," p. 61.
 Ewing, J. A. and Humfrey, J. C. W. 1902 *Philos. Trans. A*, **200**, 241.
 Greenland, K. M. 1937 *Proc. Roy. Soc. A*, **163**, 29.
 Greninger, A. B. and Mooradian, V. G. 1938 *Tech. Publ. Amer. Inst. Min. Engrs*, no. 867.
 Kanter, J. J. 1937 *Amer. Inst. min. metall. Engrs*, Iron Steel Met. Meeting, Oct. 1937.
 Orowan, E. 1934 *Z. Phys.* **89**, 605.
 Schäfer, K. 1933 *Naturwissenschaften*, **10**, 207.
 Taylor, G. I. 1928 *Proc. Roy. Soc. A*, **118**, 1.
 — 1934a *Proc. Roy. Soc. A*, **145**, 362.
 — 1934b *Proc. Roy. Soc. A*, **145**, 388.
 Taylor, G. I. and Elam, C. F. 1926 *Proc. Roy. Soc. A*, **112**, 337.
 Tsien, L. C. and Chow, Y. S. 1937 *Proc. Roy. Soc. A*, **163**, 19.

Relaxation methods applied to engineering problems

III. Problems involving two independent variables

BY D. G. CHRISTOPHERSON, B.A., AND R. V. SOUTHWELL, F.R.S.

(Received 8 June 1938)

1. Problems already treated in this series (Bradfield and Southwell 1937; Black and Southwell 1938) have been concerned both with systems of restricted freedom (stress determination in framed structures; current partition in electrical networks; adjustment of errors of observation) and with continuous systems governed by equations in one variable (transversely loaded beams). For the latter two lines of attack have been successful: either the governing equation can be replaced by an approximate equation in finite differences which is soluble, or by an application of Relaxation Methods it can be satisfied as it stands, not at all but at a finite

number of "sections", i.e. for a finite number of values of the independent variable.* Either procedure leads to tabulated values of the function investigated, for values of the independent variable separated by small intervals. Solutions thus presented are of more immediate value than the mathematical expressions given by orthodox analysis, which may require much labour to be expended in numerical computation before their significance can be appreciated. Tested in certain cases which can be solved exactly, both methods seem capable of giving more than sufficient accuracy for practical purposes.

Methods which permit a comparable treatment of problems in two dimensions will have still greater value, because here the power of orthodox methods is more restricted and in general their solutions are more difficult to interpret numerically. We must expect to encounter greater difficulties, seeing that terminal conditions are now replaced by conditions relating to every point of some geometrical boundary; and it is significant that much attention has been devoted to the development of *experimental* methods, based on analogies whereby equations occurring in one branch of physics can be solved by experiments made in another, because mathematically the same equation describes the phenomena of both. The "membrane analogue" of Prandtl (1903), utilized in the "soap-film method" of Taylor and Griffith (1917) for solving de Saint-Venant's problems of torsion and flexure (Southwell 1936, Appendix to Chapter XI), and the electrical method used by Relf (1924) and others to determine hydrodynamic streamlines, may be cited as examples.† Mechanical equation-solvers of the kinds developed in recent years by Bush, Hartree, Mallock and others do not seem likely to provide much assistance in this field.

Of theoretical methods concerned with equations in two variables three may be cited here: (1) a very ingenious solution of Laplace's equation by a synthesis of solutions having singularities at the boundary, developed by Bairstow and Berry (1919) and applied to the torsion problem by Bairstow and Pippard (1922); (2) an extension of the same method to the biharmonic equation, propounded by Bairstow, Cave and Lang (1923); (3) an approximate numerical method, applicable to both of the equations mentioned, which has been proposed by Thom (1928).

* This method of approximating to an exact solution has been termed by Frazer (1937) the "method of collocation". It assumed without proof (but with plausibility as regards problems of flexure and the like) that if the governing equation is satisfied exactly at several points in a range, it will be satisfied *approximately* at intermediate points.

† Recently the electrical method has been employed to effect the operation of conformal transformation (Bradfield, Hooker and Southwell 1937).

2. This paper is concerned with Poisson's equation

$$\nabla^2 w \equiv \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right] w = Z, \quad (1)$$

including Laplace's plane-harmonic equation as a special case ($w=0$). The method which it proposes is in effect a combination of the two methods outlined in § 1: that is to say, in the first place (1) is replaced by an approximate equation involving finite differences, and in the second this modified equation is satisfied *approximately* by a relaxation process. The finite-difference equation is a generalization (making for better approximation) of that proposed by Thom: the idea of "systematic relaxation" enables it to be applied with greater flexibility, and to a wider range of problems. Of necessity (since two steps, instead of one, separate the exact from the approximate solution) more labour must be expended, in order to obtain results of a given accuracy, than in the simpler problems which have been treated previously; but the labour is not prohibitive, and it is offset by the power of the method to deal with problems which are difficult or even intractable by exact analysis.

3. The arrangement of the paper may be summarized as follows:—In §§ 4–13 finite-difference equations corresponding with (1) are derived and their approximation considered; by means of a particular physical interpretation (namely, that in which w and Z stand for the transverse displacement and loading of a stretched and initially flat membrane) they are brought within the scope of the relaxation method, and a technique for numerical computation is outlined. In §§ 14–18 this technique is applied to a straightforward plane-potential problem,—the solution of Saint-Venant's "torsion problem" for a triangular shaft.

The membrane analogue of the torsion problem, originally suggested by Prandtl (§ 1), is already well known. In that analogue the membrane may either be unloaded, when it serves to determine the "associated torsion function" ψ , which is plane-harmonic, or it may be loaded by a uniform pressure, when it serves to determine the "shear stress-function" Ψ . The latter is governed by the equation

$$\nabla^2 \Psi = -2, \quad (2)$$

—a special case of equation (1). Its boundary condition is more convenient than that imposed on ψ —especially when the cross-section is "multiply-connected"; and whereas to orthodox analysis Poisson's equation presents

a harder problem than Laplace's, when relaxation methods are employed there is nothing to choose between the two equations. The point is illustrated in §§ 19–27, where the torsion problem is solved for a triangular section pierced by three axial holes: this second example reveals more clearly the value of the method, and serves to illustrate additional points of detail.

In §§ 28–30 the method is applied to a still harder problem,—determination of the shear-stress distribution in a shaft which has been twisted sufficiently to overstrain the material. Orthodox methods would appear to be incapable of treating this problem in *plasticity*, for the reason that it is not possible *a priori* to define the region within which the material is overstrained; and accordingly in the past recourse has been had to a very elegant, but difficult, experimental technique (§ 29). The methods of this paper lead without difficulty to a solution.

As a final example, in §§ 31–35 lines of flux are determined for a two-dimensional magnetic field containing a triangular prism of iron. This problem being also intractable by orthodox methods, it was investigated by Hele-Shaw and Hay (1900) with the aid of a very beautiful “stream-line apparatus”. Relaxation methods are again successful in this instance.

I. GENERAL DESCRIPTION OF THE APPROXIMATE METHOD

Derivation of the finite-difference equation

4. Our aim is to obtain an approximate solution of (1), Z being specified; that is to say, we require to determine w (approximately) at selected points within a specified region. The number of these points may be large, but in the nature of the case it must be finite, and thus the word “solution” has a different significance here from what it has in orthodox mathematics, where it stands for an expression in functional form from which the value of w could be calculated for *every* point in the region. Here the significance is less precise, because an indefinite number of functions will have the specified values at the selected points and be single-valued elsewhere, and any one will have the same claim as any other to be regarded as the required solution. Adopting a standpoint which is the basis of the theory of interpolation (Whittaker and Robinson 1926, § 8) we may say that any “smooth” function of the variables x, y which takes the specified values at the selected points will be, for practical purposes, a satisfactory solution; and for simplicity we choose a *polynomial* function, which can always be formulated.*

* See Appendix.

Let \bar{w} stand for the exact solution, \bar{Z} for the specified function with which it is related by (1), w for the polynomial approximation, and Z for the function which would be required on the right-hand side of (1) in order to make w an exact solution of that equation. Evidently Z , like w , will be a polynomial function of x and y , and so cannot (in general) be identified with \bar{Z} . On the other hand, Z can be made equal to \bar{Z} at each of the selected points, and then, if \bar{Z} is a "smooth" function, Z will be a close approximation everywhere.

5. Because Z and w are polynomials their differential coefficients, of all orders, will be finite everywhere: therefore it would seem that close accuracy is not to be expected in cases where either \bar{Z} or \bar{w} is characterized by singularities. But because equation (1) is linear, solutions can be superposed; moreover, exact solutions of (1) are known for most types of singularity which will be presented. Accordingly we can usually dispose of singularities in advance, and so reduce our problem to one in which the polynomial approximation is sufficient. The point will not arise in any of the problems treated in this paper.

6. We now proceed (cf. § 2) to replace equation (1) by a finite-difference relation between w and Z , Z being defined as a polynomial function agreeing at the selected points with the specified function \bar{Z} .

Let x, y be replaced by polar co-ordinates r, θ , measured with respect to an origin at 0. Then in the altered co-ordinates w (being a polynomial in x and y) can be expressed in the form

$$w = A_0(r) + \Sigma_n [A_n(r) \cos n\theta + B_n(r) \sin n\theta], \quad (3)$$

where the n 's are integral and $A_0(r), A_n(r), B_n(r)$ are polynomial functions of r ; and according to (1) the corresponding expression for Z is

$$\begin{aligned} Z &= \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} \right] w, \\ &= A_0'' + \frac{A_0'}{r} + \Sigma_n \left[\left(A_n'' + \frac{1}{r} A_n' - \frac{n^2}{r^2} A_n \right) \cos n\theta \right. \\ &\quad \left. + \left(B_n'' + \frac{1}{r} B_n' - \frac{n^2}{r^2} B_n \right) \sin n\theta \right], \end{aligned} \quad (3A)$$

dashes denoting differentiations with respect to r .

Now let $\Sigma_{a, N}(w)$ stand for the sum of the values assumed by w at N points

equally spaced on the circle $r = a$, so that their angular positions are $\theta = 0, \alpha, \dots, (N-1)\alpha$, where $\alpha = 2\pi/N$. Then according to (3)

$$\begin{aligned}\Sigma_{a,N}(w) &= NA_0(a) + \Sigma_n[A_n(a)\{1 + \cos n\alpha + \dots + \cos(N-1)n\alpha\} \\ &\quad + B_n(a)\{0 + \sin n\alpha + \dots + \sin(N-1)n\alpha\}], \\ &= NA_0(a) + \Sigma_n\left[\frac{\sin \frac{1}{2}Nn\alpha}{\sin \frac{1}{2}n\alpha}\{A_n(a)\cos \frac{1}{2}(N-1)n\alpha\right. \\ &\quad \left.+ B_n(a)\sin \frac{1}{2}(N-1)n\alpha\}\right],\end{aligned}$$

by known formulae in trigonometry. Since $N\alpha = 2\pi$ and n is integral, the quantity in square brackets will be zero unless $\sin \frac{1}{2}n\alpha$ vanishes,—that is, unless $n\alpha$ is some integral multiple of 2π : in that event n must be an integral multiple of N , and the quantity takes the value $NA_n(a)$. Thus we see that

$$\frac{1}{N}\Sigma_{a,N}(w) = A_0(a) + A_N(a) + A_{2N}(a) + \dots, \quad (4)$$

N being the number of points at which values of w are taken.

7. Now it is easy to show (w being polynomial in x and y) that $A_n(r)$, $B_n(r)$, in (3), can involve no power of r lower than the n th and must in fact have the form $r^n \times (\text{polynomial in } r^2)$. Therefore the neglected terms will be of order a^N at least when we replace (4) by the approximate equation

$$\frac{1}{N}\Sigma_{a,N}(w) = A_0(a). \quad (5)$$

Moreover, $\cos n\theta$ and $\sin n\theta$, in (3 A), will be associated with r^n and higher powers of r , so that only A_0 will contribute to the value of Z at the origin.

Again, $A_0(r)$ must be polynomial in r^2 , so that we may write

$$A_0(r) = c_0 + c_2r^2 + c_4r^4 + \dots \text{ etc.}, \quad (6)$$

where c_0, c_2, \dots , etc., have constant values. Therefore at the origin ($r=0$)

$$Z = 4c_2,$$

and we can show that $\nabla^2 Z = 4.16c_4$,

$$\nabla^4 Z = 2^2.4^2.6^2c_6,$$

\dots , etc.,

if we apply the same reasoning to $Z, \nabla^2 Z, \dots$, etc., since these, like w , are polynomials in x and y . Finally, at the origin we have from (3) and (6)

$$w = c_0.$$

Substituting these values for c_0, c_2, \dots , etc., in (6), we find that

$$A_0(r) = w_0 + \frac{r^2}{4} Z_0 + \frac{r^4}{64} (\nabla^2 Z)_0 + \frac{r^6}{2^2 \cdot 4^2 \cdot 6^2} (\nabla^4 Z)_0 + \dots,$$

in which the suffixes attached to w, Z, \dots , etc., denote that the values are to be taken at the origin; and the approximate relation (5) may thus be written in the form

$$\frac{1}{N} \Sigma_{a,N}(w) - w_0 = \frac{a^2}{4} Z_0 + \frac{a^4}{64} (\nabla^2 Z)_0 + \frac{a^6}{2^2 \cdot 4^2 \cdot 6^2} (\nabla^4 Z)_0 + \dots \quad (7)$$

8. Already, in substituting (5) for (4), we have neglected terms of order a^N : therefore it is useless to retain terms of this or higher order on the right-hand side of (7), but within this restriction we may proceed to find expressions for $(\nabla^2 Z)_0, \dots$, etc. (We cannot calculate these quantities directly, for the reason that Z cannot be identified with the specified function \bar{Z} , except at the selected points.)

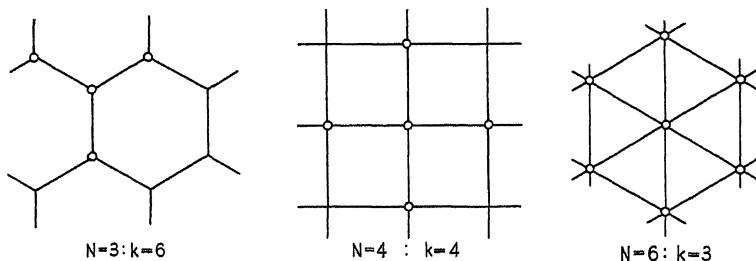


FIG. 1

The underlying idea of this investigation is that the selected points will lie on some regular lattice, or "net", such that the summation $\Sigma_{a,N}$ can be effected at every point. We shall be given values of $\bar{Z}(=Z)$ at the selected points, and by similar reasoning we can derive relations similar to (7), viz.

$$\left. \begin{aligned} \frac{1}{N} \Sigma_{a,N}(Z) - Z_0 &= \frac{a^2}{4} (\nabla^2 Z)_0 + \frac{a^4}{64} (\nabla^4 Z)_0 + \dots, \\ \frac{1}{N} \Sigma_{a,N}(\nabla^2 Z) - (\nabla^2 Z)_0 &= \frac{a^2}{4} (\nabla^4 Z)_0 + \frac{a^4}{64} (\nabla^6 Z)_0 + \dots, \\ &\dots, \text{ etc.} \end{aligned} \right\} \quad (8)$$

Lattices permitting the summation $\Sigma_{a,N}$ at every point can be obtained (fig. 1) by giving N any of the values 3, 4 or 6, but higher values are not

admissible.* Therefore it is useless to retain in (7) terms which involve a^6 or higher powers of a , but *with an error of order a^N (at least) in each case* we have, from (7) combined with (8),

$$\text{For } N = 3 \text{ or } 4: \quad \frac{1}{N} \Sigma_{a,N}(w) - w = \frac{a^2}{4} Z; \quad (9)$$

$$\begin{aligned} \text{For } N = 6: \quad \frac{1}{N} \Sigma_{a,N}(w) - w &= \frac{a^2}{4} \left[Z + \frac{1}{4} \left\{ \frac{1}{N} \Sigma_{a,N}(Z) - Z \right\} \right], \\ &= \frac{a^2}{16} \left[3Z + \frac{1}{N} \Sigma_{a,N}(Z) \right]. \end{aligned} \quad (10)$$

Suffixes have been deleted as no longer necessary.

These are our finite-difference equations, of which (9) can be used whether $N = 3, 4$ or 6 , and (10)—for closer approximation—only when $N = 6$. The first step towards a solution of (1) will be to substitute one or other of them for that equation: the next will be to satisfy the approximate relation at each of the selected points.

A mechanical analogy

9. Thom (1928), by different reasoning which envisaged the use of a square mesh ($N = 4$), deduced a finite-difference equation identical with (9)† and showed that the neglected terms were of order a^4 at least, thus confirming our result. In solving his equation he made (in the nature of the case) no use of relaxation ideas. These are not essential to the solution of (9) or (10), but they have value as saving labour.

The nomenclature of the relaxation method is appropriate to the mechanical systems for which they were originally devised: therefore it will be convenient to interpret equation (1) as relating to the transverse deflection of a membrane, initially flat, which is stretched by a uniform tension T . If the membrane is loaded by a small transverse pressure of intensity $-TZ$, then w as governed by (1) will measure the transverse displacement, and the boundary condition will impose specified values either on w or on the line intensity of transverse force $\left(T \frac{\partial w}{\partial \nu}\right)$. Thus every plane-potential problem has a membrane analogue, and Prandtl's analogue (§ 1) is merely a particular example.

The finite-difference equations (9) and (10) can be interpreted similarly.

* The condition requires that $2\pi/N$ shall be the internal angle of a regular polygon: i.e. (fig. 1) that $N - 2 = \frac{4}{k - 2}$, where N and k ($k \leq 3$) are integral.

† Cf. equation (12), § 10.

A flat net made from inextensible strings, and having meshes of the form of a regular polygon (fig. 1), can remain in equilibrium with every string strained to the same tension T . If now the nodes are given transverse displacements measured by w , the transverse force exerted on any one node by the strings which radiate from it will be measured by $\frac{T}{a}\{\Sigma_{a,N}(w) - Nw\}$: therefore according to (9) and (10) the node will be in equilibrium when loaded by a transverse force (in the direction of w) given by

$$\left. \begin{aligned} \bar{F} &= -\frac{1}{4}NTaZ, \text{ in the first approximation } (N = 3 \text{ or } 4) \\ &= -\frac{1}{16}NTa\left[3Z + \frac{1}{N}\Sigma_{a,N}(Z)\right], \\ &\text{in the second approximation (valid when } N=6\text{).} \end{aligned} \right\} \quad (11)$$

Our replacement of (1) by (9) or (10) is in effect a replacement of the continuous and continuously loaded membrane by a net loaded at its nodes, and the accuracy of our results may be expected to improve as the size (a) of the mesh is reduced.

Outline of the relaxation procedure

10. Thom, in dealing with Laplace's equation ($Z=0$), begins by attaching plausible values to w at nodal points of his (square-mesh) net, and thereafter systematically modifies these values by applying the equation

$$w = \frac{1}{4}\Sigma_{a,N}(w) \quad (12)$$

to every point in turn: in time the modifications become inappreciable, and then the modification process is stopped. Equation (12) being identical with (9) when $Z = 0$, $N = 4$, his process may be regarded as an application of relaxation methods whereby at every step he (temporarily) relieves some particular constraint of the residual force previously sustained by it. We now interpret from this standpoint the more general equations (9) and (10).

11. Visualizing constraints which operate to control the displacements (w) of nodal points in a chosen net, we write

$$\bar{F} = F + \bar{F}, \quad (13)$$

for the "residual" force (in the direction of w) which acts on any one nodal constraint. Of the two parts into which \bar{F} is here divided,

\bar{F} , as given by (11), is the external force at the node in question, and

F , given by $N\frac{T}{a}\left\{\frac{1}{N}\Sigma_{a,N}(w) - w\right\}$, is the resultant force exerted, on account

of the displacement, by the N strings which radiate from this node.

For equilibrium all the \bar{F} 's must vanish, so as to satisfy (9) or (10) at every one of the selected points. They can be calculated for any trial solution and "liquidated" by changes made systematically in the w 's. For this purpose we shall want to know the effects on the \bar{F} 's of an isolated displacement Δw imposed at any one node. The \bar{F} 's are not affected, and from the expression for F we have

At the node which is moved:

$$\Delta \bar{F} = \Delta F = -N \frac{T}{a} \Delta w,$$

At each of the N surrounding nodes:

$$\Delta \bar{F} = \Delta F = +\frac{T}{a} \Delta w.$$

From these results and from (11) it is evident that the final displacements (w) will be independent of the value of NT/a : therefore we now (for simplicity) take that quantity as equal to unity. On this understanding, when a unit displacement is imposed on any one node, the residual forces will be altered

by -1 at the node which is moved,

by $+\frac{1}{N}$ at each of the N surrounding nodes,

and the expressions (11) for the initial force become

$\bar{F} = -\frac{1}{4} a^2 Z$, in the first approximation ($N = 3$ or 4),

$= -\frac{1}{16} a^2 \left[3Z + \frac{1}{N} \Sigma_{a,N}(Z) \right]$, in the second approximation ($N = 6$). (14)

Elimination of dimensional factors

12. Equation (1) is "dimensional", in that x and y are co-ordinates measured in some particular units. But if L is some representative dimension of the region within which w is to be determined, we may write

$$x = L\xi, \quad y = L\eta, \tag{i}$$

and

$$w = Z_0 L^2 W(\xi, \eta), \quad Z = Z_0 \zeta, \tag{ii}$$

where Z_0 stands for the value of Z at some particular point in the region; and on substituting these expressions in (1) we obtain

$$\left[\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} \right] W = \zeta, \tag{iii}$$

—an equation similar in form to (1), but in which ξ , η , W , ζ are purely numerical quantities. Its solution will define the function $W(\xi, \eta)$, and then we shall have in (ii) a solution of (1) which applies to all regions of a particular shape and similarly “loaded”, without restriction on their size or on the absolute intensity of the “loading”.

In §§ 6–8, similarly, we may express r and a in terms of the dimensional parameter L by writing

$$r = L\rho, \quad a = L\alpha. \quad (\text{iv})$$

Then, in place of (9) and (10), we shall have finite-difference relations—namely,

$$\text{For } N = 3 \text{ or } 4: \quad \frac{1}{N} \Sigma_{a,N}(W) - W = \frac{\alpha^2}{4} \zeta, \quad (\text{v})$$

$$\text{For } N = 6: \quad \frac{1}{N} \Sigma_{a,N}(W) - W = \frac{\alpha^2}{16} \left[3\zeta + \frac{1}{N} \Sigma_{a,N}(\zeta) \right], \quad (\text{vi})$$

—which involve only numerical quantities (N , W , α , ζ), and are thus appropriate to numerical computation.

It will not in fact be necessary to change the symbols in this way, provided that w , Z , x , y , a in our equations are given the non-dimensional meanings which in this section have been attached to W , ζ , ξ , η , α . In what follows this is to be understood, so that a , in (9), (10) and (14), will henceforth define the length of one side of a “mesh”, expressed as a fraction of L .

The use of “successive nets”

13. The labour as well as the accuracy of a solution will increase as the size of the mesh is reduced. Therefore in practice it will be best to start with a net of large mesh, liquidate the “residual force” at every nodal point within the boundary, and from this first approximation derive a trial solution for a net of smaller mesh. A suitable procedure is indicated in fig. 2. The first net is shown in bold lines, and its nodes by black circles: having values of w at these points we can from (9), with $N = 3$, deduce a value for w at the centroid of each mesh in terms of its values at the three corners. The centroids (open circles), joined by broken lines with one another and by fine lines with the nodal points of the original net, become the nodes of a new triangular net; and for these we have values available as a trial solution, to be corrected by relaxation methods as described in § 11. The process can be repeated indefinitely, and at each repetition the side of a mesh will be reduced in the ratio $1 : \sqrt{3}$. If in each relaxation process we give to N the value 6, then every

successive approximation will be as accurate as the scale of its net permits; and since the omitted terms are of order α^N at least (§ 8), a decrease in the errors of the order of $1/27$ may be expected.

We now apply these ideas to a particular example, leaving for subsequent discussion the special problems which may be presented at the boundary.

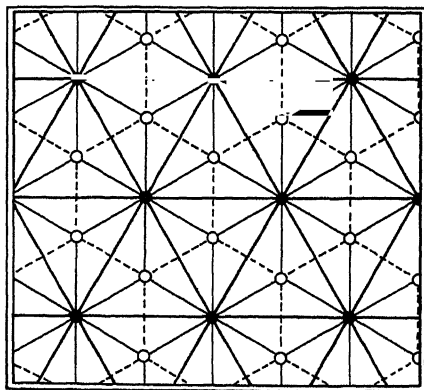


FIG. 2

II. A STANDARD EXAMPLE IN PLANE-POTENTIAL THEORY: THE TORSION PROBLEM FOR A TRIANGULAR SHAFT

14. To solve the torsion problem for a shaft of given section we have (Love 1927, § 217) to find a plane-harmonic function satisfying the condition

$$\psi = \frac{1}{2}(x^2 + y^2) + \text{const.} \quad (15)$$

at every point on its boundary. As a first example we shall solve this problem for the case of an equilateral triangle.

The known solution (for an origin at the centroid) is

$$\psi = (x^3 - 3xy^2)/6h + \text{const.} \quad (16)$$

when the boundary is given by

$$(x+h)(x-y\sqrt{3}-2h)(x+y\sqrt{3}-2h) = 0,$$

so that the height of the triangle is $3h$.* If the constant in (16) is given the value $h^2/6$, then on the side $(x+h=0)$

$$\psi = \frac{1}{2}y^2,$$

so that (15) is satisfied, and the same is true of the other two sides. Thus ψ is a plane-harmonic function which along any side of the triangle has values

* Cf. Bradfield, Hooker and Southwell (1937), § 6.

proportional to the square of the distance from the middle point. The constant of proportionality is immaterial as concerns the potential problem: to avoid fractional values in computation we shall assume that $\psi = 2025$ ($= 45^2$) at the corners.

15. \bar{Z} being zero, we have $\bar{F} = 0$ according to (14), i.e.

$$\bar{F} = F = \frac{1}{N} \sum_{a, N} (w) - w, \quad (17)$$

w denoting the approximate solution for ψ , and \bar{F} the residual force (which must be made to vanish at every node within the triangular boundary). We start with a net dividing each side of the triangle into three (fig. 3) and having one internal node O : there, according to (17) with $N = 6$, \bar{F} will vanish if w has the value 225 ($= 15^2$), as at the six surrounding nodes.

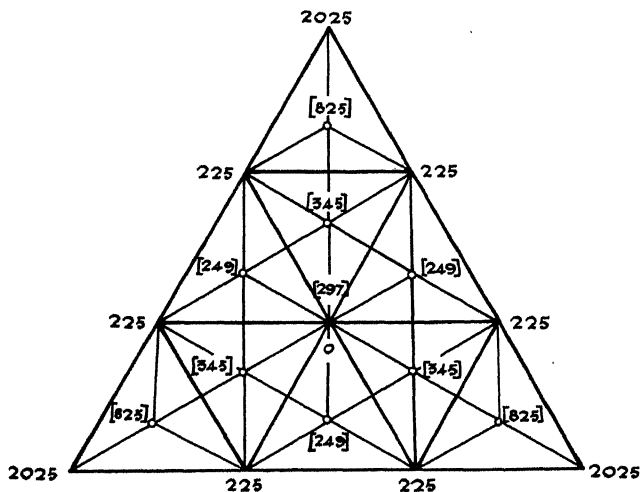


FIG. 3

Next (§ 13) we calculate values for the centroids (open circles) of the nine meshes, using (17) with $N = 3$; then we correct this trial solution for the smaller net (fine lines in fig. 3). In this particular instance successive corrections were observed to decrease in geometrical progression, and on that account \bar{F} could be made to vanish everywhere, and exactly,* for values of w which are shown in fig. 3. These gave, before, a trial solution for the next net (fig. 4a).

16. Thereafter the calculations followed standard lines. From the displacements given on the left of fig. 4a (by means of (17) with $N = 6$ where

* N being taken as 6 where this is possible, otherwise as 3.

ruled net: therefore it is convenient to keep them on a transparent sheet laid over the net, using a soft pencil in order that figures may be erased from time to time after their totals have been recorded. Convenient material for this purpose is marketed by Messrs Kodak, Ltd., under the name "Kodatracer". It is durable and highly transparent, so that erasures can be made almost indefinitely.

In this example nine operations sufficed to reduce the numerical value of the largest residual force from about 54 to about 3: subsequently the residual forces were multiplied by 100 and again liquidated until, none having a value as large as 10, the accuracy of the approximation was deemed sufficient. A summary of the relaxation process up to this stage is given in fig. 4c, numbers on the right of each node relating to residual forces and numbers on the left to additional displacements imposed at that node. Adding the initial displacements (fig. 4a) we obtain the total values distinguished by underlining: if displacements are wanted only to the nearest whole number, these constitute a final approximation for this particular size of net.

17. Actually the arrangement of the residual forces (which are all either zero or $\pm 1/3$) is such that they can be liquidated *completely* by the imposition of a further "displacement" on each node equal (numerically) to the residual force, so that (e.g.) the number 758 becomes $758\frac{1}{3}$ and the number 292 becomes $291\frac{2}{3}$. *Thus adjusted, the numbers are in fact correct.* This was not noticed until the relaxation procedure had been effected for two further nets, the last having meshes of side $L/27$, where L is the length of a side of the triangular boundary; but it is easy to see that in this example complete liquidation of the residual forces will always give an exact result when the net is such that even close to the boundary each nodal point is surrounded by 6 others of which the displacements are defined: for the exact expression for ψ (§ 14) is cubic in x and y , so that in (3) the only values assumed by n are 1, 2 and 3; and in consequence, when $N = 6$, only $A_0(a)$ appears on the right-hand side of (4), so that (5) is *exact*. When $N = 3$ the relation (5) is only an approximation to (4), obtained by suppression of a term $A_3(a)$ on the right of that equation.

18. Reverting to general aspects we observe, first, that much time is saved in the relaxation method by its concentration on *residual* forces and *corrections* to the displacements, as compared with Thom's iterative process which at every stage is concerned with *total* displacements. As the mesh becomes smaller, smaller quantities are involved: in fact, to avoid fractions it is usually convenient to multiply them in the later stages.

Secondly, since the operations employed in the relaxation process included "group displacements", it should be observed that to deduce these from (14) is an extremely simple matter. When one node only is displaced, each of the "strings" which radiate from it must (by symmetry) make the same contribution to the change in the residual force: therefore we may think of each string as *transferring* a force $1/N$ from the end which is displaced to the end which is fixed; and thus, when a "block" of nodes is displaced, we have only to consider the total residual force on that block, and the number of strings affected, in order to determine a displacement which will bring the total force to zero. The "blocks" displaced in this example were hexagons surrounding the central point: two such hexagons were moved.

III. AN EXAMPLE INVOLVING POISSON'S EQUATION: THE TORSION PROBLEM FOR A MULTIPLY-CONNECTED SECTION

19. Our second example (fig. 5) is a section containing three holes, therefore three internal boundaries. Its symmetry permits the calculations to be described more shortly than would otherwise be possible, but it has no feature tending to restrict their generality.

Multiply-connected sections are most conveniently treated in terms of the "shear-stress function" Ψ (§ 3). When ψ is given by (15), the function

$$\Psi = \psi - \frac{1}{2}(x^2 + y^2) \quad (18)$$

satisfies the equation

$$\nabla^2 \Psi = -2 \quad (2 \text{ bis})$$

and has a constant value on the boundary. The constant is immaterial as regards "solid" sections, and to make the plane-potential problem definite we may say that it is zero. We might have worked Example I in terms of Ψ .

When the section is "multiply-connected" Ψ can again be made zero on the outer boundary, but then at an internal boundary its value is determined by a further requirement—namely, that ϕ , the plane-harmonic function which is conjugate to ψ , must (since it measures the axial displacement due to torsion) be single-valued at every point in the section. Therefore the integral

$$\int \frac{\partial \phi}{\partial s} ds = \int \frac{\partial \psi}{\partial \nu} ds,$$

taken round any one boundary, must be zero, ν denoting the normal to the boundary drawn outwards; and it follows from (18) that

$$\int \frac{\partial \Psi}{\partial \nu} ds = -\frac{1}{2} \int \frac{\partial}{\partial \nu} (x^2 + y^2) ds = -2A, \text{ by Green's transformation,} \quad (19)$$

A denoting the area contained within that boundary.

In the membrane analogue (§ 9) Ψ as governed by (2) is the deflection produced by a uniform pressure acting at every point, and the foregoing result shows that when a section has internal boundaries we must think of each as the edge of a light rigid plate, permitted to move perpendicularly to its own plane, to which the membrane is attached and on which the uniform pressure is operative. Our approximation, in effect, replaces the membrane by a net (§ 9), and the condition that each plate must "float" freely in the manner just described would be difficult to satisfy directly. Instead, we now consider how the required solution may be derived by a synthesis of partial solutions.

20. Suppose first that the plates are prevented from moving, so that $\Psi = 0$ both at the external and internal boundaries: let the uniform pressure be applied, and let Ψ_0 stand for the displacement of the membrane under these conditions. Next, after removal of the pressure let one plate be given a unit displacement while the others are kept fixed, so that $\Psi = 1$ at one internal boundary but vanishes at all the other boundaries: let Ψ_1 stand for the displacement of the membrane under these conditions (Ψ_1 will be a plane-harmonic function of x and y). Lastly, let the other two internal boundaries be moved in turn, and let Ψ_2, Ψ_3 (also plane-harmonic functions of x and y) stand for the resulting displacements in the two cases.

Then a solution defined by

$$\Psi = \Psi_0 + k_1 \Psi_1 + k_2 \Psi_2 + k_3 \Psi_3 \quad (20)$$

(k_1, k_2, k_3 being constants) will evidently satisfy the governing equation (2) and will have a constant value on every boundary. Moreover, if A_1, A_2, A_3 stand for the areas of the three plates, the condition of "floating" will be satisfied for A_1 if

$$-2A_1 = \int_1 \frac{\partial \Psi}{\partial \nu} ds = \int_1 \frac{\partial \Psi_0}{\partial \nu} ds + k_1 \int_1 \frac{\partial \Psi_1}{\partial \nu} ds + k_2 \int_1 \frac{\partial \Psi_2}{\partial \nu} ds + k_3 \int_1 \frac{\partial \Psi_3}{\partial \nu} ds \quad (21)$$

when each integral is taken round the whole boundary of A_1 . Two similar equations will ensure the "floating" of A_2 and A_3 , and from the three we can deduce values of k_1, k_2, k_3 for which Ψ as given by (20) satisfies all of the imposed conditions. Any number of internal boundaries can be treated in this manner.

In the present example (fig. 5) the work can be shortened on account of symmetry. We know that all three plates must be displaced by the same amount; so the three partial solutions Ψ_1, Ψ_2, Ψ_3 can be replaced by one, in

which all three are given a unit displacement. Ψ_0 having the same significance as before, a single equation

$$-2A = \int \frac{\partial \Psi_0}{\partial \nu} ds + k \int \frac{\partial \Psi_1}{\partial \nu} ds \quad (22)$$

(in which A stands for the area of one plate, and each integral is to be taken round the whole of one internal boundary) will suffice to determine a value of k such that

$$\Psi = \Psi_0 + k\Psi_1 \quad (23)$$

satisfies all of the imposed conditions.

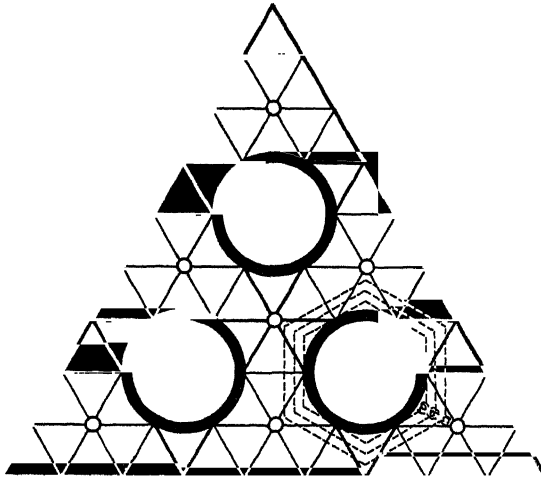


FIG. 5

Elimination of dimensional factors

21. In conformity with § 15 we denote the approximate solutions for Ψ_0 , Ψ_1 by w_0 , w_1 respectively. The first, w_0 , has to vanish on every boundary and to satisfy (9) and (10)—which are here identical—at every nodal point in the field.

We can interpret (2) “non-dimensionally” in the manner of § 12 by regarding x and y as coordinates expressed in terms of L , the length of a side of the equilateral triangle. Since $Z = -2$ in equation (1), if we take Z_0 as unity then Z will have the same value interpreted non-dimensionally, so that both of the formulae (14) for \bar{F} reduce to

$$\bar{F} = \frac{1}{2} a^2, \quad (24)$$

a denoting the side of a mesh expressed as a fraction of L . In accordance with §12 our numerical values of w_0 , multiplied by L^2 , will constitute the approximate solution for Ψ_0 .

The other formulae (14), for the changes caused by a unit displacement, will apply without alteration; and both (9) and (10) will give

$$\bar{F} = \frac{1}{2}a^2 + \frac{1}{N}\sum_{a,N}(w_0) - w_0 = 0. \quad (25)$$

First partial solution

22. On account of the internal (circular) boundaries we are not obliged to start from the coarse net ($a=1/3$) shown by bold lines in fig. 5, but can proceed at once to the smaller mesh shown by fine lines ($a=1/9$). Each internal boundary passes through six nodal points at which w_0 is zero. Near to each corner, and to the middle of each side of the triangle, is a node surrounded by six at which w_0 is zero, and the same is true of the central point. Accordingly at all seven of these nodes (open circles in fig. 5) we have, from (25) with $a = 1/9$,

$$w_0 = \frac{1}{162}.$$

At other nodes w_0 is zero.

On this understanding figures on the left of fig. 6 give starting values of the "displacements", similar figures to the right of the centre line give (to one place of decimals) the corresponding initial forces. The ruled net is the finer net of fig. 5, and values of w_0 at the centroids of its meshes (open circles) have been calculated from (25) with $N = 3$, $a = 1/9 \sqrt{3}$. To avoid fractions in the subsequent computations *all figures have been multiplied by 43740* ($= 60 \times 27^2$), and allowance for this multiplication will have to be made in the solution finally accepted.

Satisfaction of the boundary conditions

23. In calculating values of the initial forces, and in their subsequent liquidation, a problem is now confronted which was not a feature of our first example. Close to a boundary (either internal or external) it may not be possible to employ the formulae (14) and (25), for the reason that surrounding points for which values of w_0 are known do not all lie at the same distance from the nodal point in question. Thus A , close to the triangular boundary in fig. 6, is connected with five equidistant points, but the sixth "string" (connecting it with the boundary) is only half as long as the others. B , close to the circular boundary, is connected with it by one short string.

We thus have need of a generalization whereby the formulae (14) and (25) can be applied to "stars" of unequal strings. This is a problem in interpolation, and exact treatment, since it entails considerable complication, is

hardly justified in a method whereby periodically (at every advance to a finer net) we calculate the errors of a trial solution and (in effect) start afresh. Therefore we now propose a simple modification of (14), based on the notion used already (§ 18) of a particular force transmitted by each several “string”.

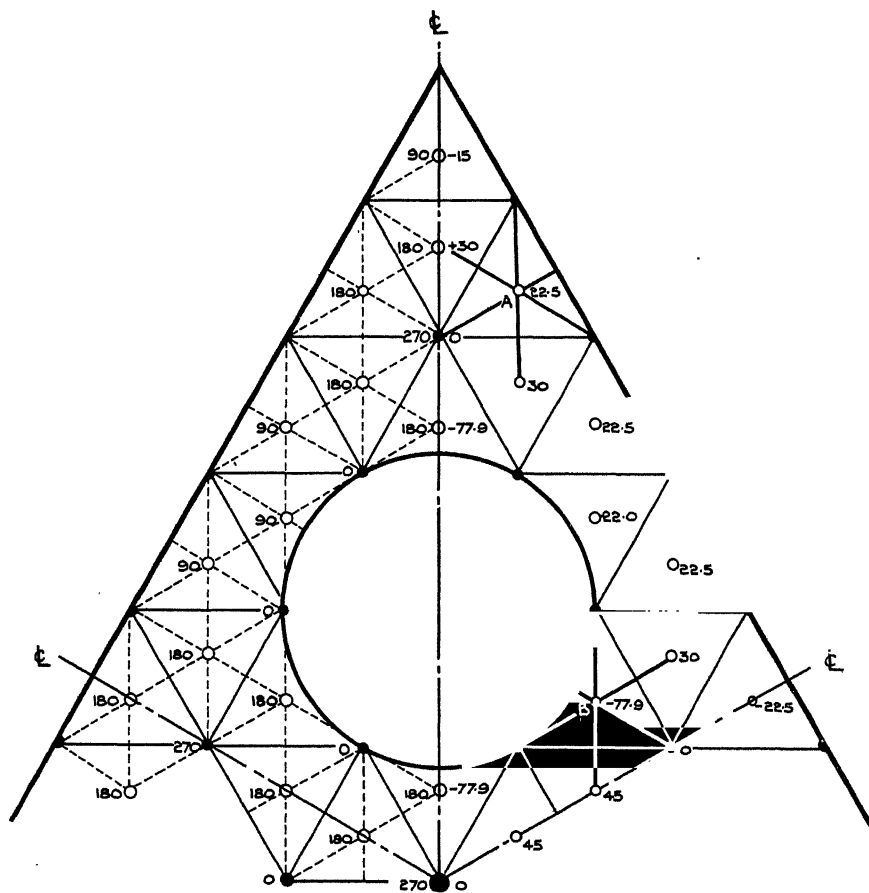


FIG. 6

In the tensioned net which was envisaged in § 9, the force exerted by a string is, for given displacements of its two ends, inversely proportional to its length. Therefore if NT/a is made unity as in § 11, and if the N strings which radiate from a displaced node have lengths $x_1 a, x_2 a, \dots, x_N a$, we have

$$\left. \begin{aligned} \Delta \bar{F} &= -\frac{1}{N} \left(\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_N} \right) \text{ at the node which is moved,} \\ &= +\frac{1}{Nx_1} + \frac{1}{Nx_2} + \dots + \frac{1}{Nx_N}, \text{ at each of the } N \text{ surrounding nodes,} \end{aligned} \right\} \quad (26)$$

in place of the formulae (14), and for calculating initial forces, in place of (17),

$$F = \frac{1}{N} \left[\Sigma_N \left(\frac{w_1}{x_1} \right) - w \Sigma_N \left(\frac{1}{x_1} \right) \right]. \quad (27)$$

24. The formulae (26) and (27) cannot, like those which they replace, be justified by an argument on the lines of §§ 6–8, because in fact the quantity within square brackets in (27) has the expression

$$\begin{aligned} \frac{1}{4} a^2 Z_0 \Sigma_N(x_1) + \Sigma_N \left[\frac{1}{x_1} \{ A_1(x_1 a) \cos \alpha + B_1(x_1 a) \sin \alpha \} \right. \\ \left. + \frac{1}{x_1} \{ A_2(x_1 a) \cos 2\alpha + B_2(x_2 a) \sin 2\alpha \} + \dots \text{etc.} \right], \end{aligned}$$

in which (cf. § 7) $A_n(r)$ and $B_n(r)$ have the form $r^n \times (\text{polynomial in } r^2)$. The x 's have different values for different strings in the "star", and thus the trigonometrical series contribute terms of order a^2 , which is also the order of the term involving Z_0 . All that can be deduced from this examination is that the term $\frac{a^2}{4} Z$, in (9) and (10), should be replaced by $\frac{a^2}{4N} Z \Sigma_N(x)$, so that in place of (24) and (25) we must use

$$\bar{F} = \frac{1}{2} \frac{a^2}{N} \Sigma_N(x) \quad (28)$$

and
$$\bar{F} = \frac{1}{2} \frac{a^2}{N} \Sigma_N(x) + \frac{1}{N} \left[\Sigma_{a,N} \left(\frac{w_0}{x} \right) - w_0 \Sigma_{a,N} \left(\frac{1}{x} \right) \right] = 0, \quad (29)$$

respectively, in relation to a star with unequal arms.

Thus at A in fig. 6 we have $a = 1/9\sqrt{3}$, $N = 6$, and all the x 's are unity except one, which is 0.5: therefore by (28) the contribution of Z to the initial force, when multiplied as in § 22, is

$$\frac{5.5}{12 \times 243} \times 43740 = 15 \times 5.5 = 82.5,$$

and according to (26) the effect of a unit displacement imposed at A will be to change the residual force

by $-\frac{7}{8}$ at A ,

by $+\frac{1}{8}$ at the five equidistant nodes,

by $+\frac{1}{3}$ at the near point of the triangular boundary.

Labour will be saved (on the whole) by using accurate values of the x 's,

because the initial forces in the next approximation (with a finer net) are thereby reduced.

Second partial solution

25. It is not necessary to describe the relaxation procedure, since (apart from the special treatment of stars with unequal arms) this followed exactly the lines of 15–16. In this first “partial problem”, Z being finite, a definite standard of approximation could be imposed—namely, the reduction of residual forces to some definite (small) percentage of their values when w_0 is zero. We carried the solution to a net for which $\alpha = 1/27\sqrt{3}$, and all residual forces were liquidated within 1 %.

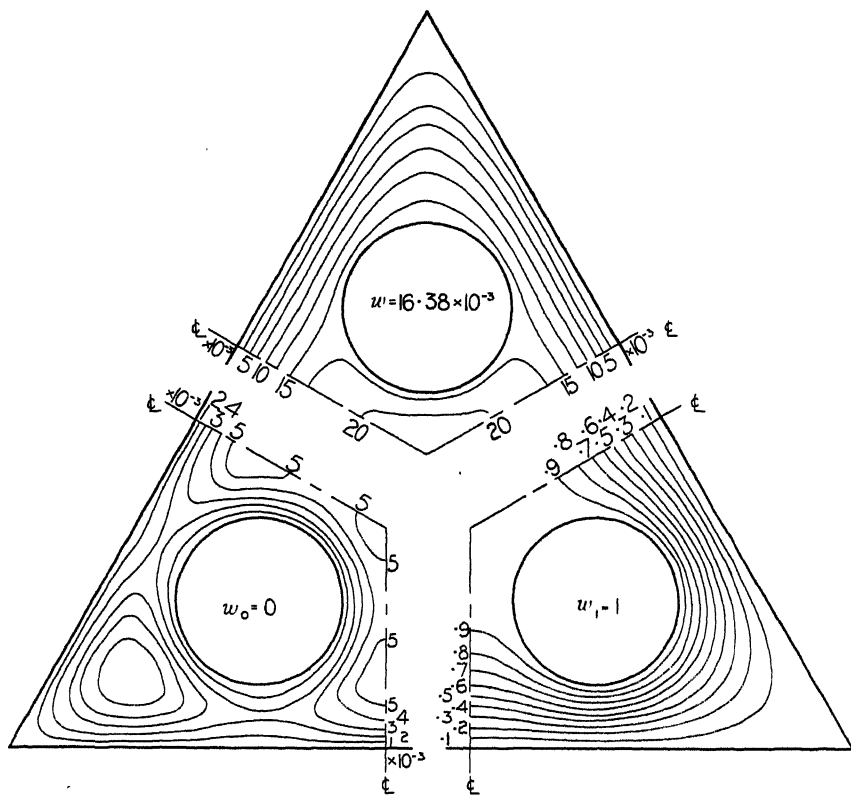


FIG. 7

The second partial solution Ψ_1 (§ 21) is a plane-harmonic function, i.e. $Z = 0$. Here too the solution was carried to a net for which $\alpha = 1/27\sqrt{3}$, and the final values of residual force were brought within the same numerical limits. Our results for both partial problems are exhibited graphically in the lower part of fig. 7, by contours of constant w numbered in the usual

manner. The multiplying factor 43740 (§ 22) was removed in these final calculations.

Estimation of the forces on internal boundaries

26. The partial solutions Ψ_0, Ψ_1 (§ 20) must be combined according to (23) so as to satisfy equation (22)—the condition that each circular plate shall “float”. We have replaced Ψ_0, Ψ_1 by approximations w_0, w_1 which have been evaluated only at nodal points: therefore we now want expressions in finite differences by which the integrals in (22) may be replaced.

In § 20 it was stated that each of the two integrals in (22) is to be taken round the whole circular boundary of one plate. But statical considerations applied to the membrane analogue (§ 19) show that the circle may be replaced by any contour which contains it but neither of the other two circles, and which lies within the external (triangular) boundary. The same conclusion may be established as a consequence of Green’s theorem, observing that Ψ_1 is plane-harmonic and Ψ_0 a solution of equation (2). It permits a choice of contour which greatly simplifies the approximate treatment of (22).

In accordance with § 21 we can (approximately) write $L^2 \times w_0, w_1$ for Ψ_0, Ψ_1 in (22); and evidently ν and s can be interpreted as multiples of L without introducing any change in “dimensions”. On that understanding we have

$$-2 \frac{A}{L^2} = \int \frac{\partial w_0}{\partial \nu} ds + k \int \frac{\partial w_1}{\partial \nu} ds \quad (30)$$

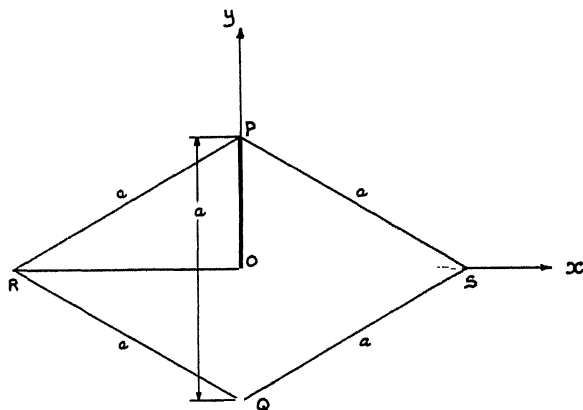


FIG. 8

as a purely numerical equation replacing (22), and in it the integrations relate to any convenient contour of area A . Now suppose that O (fig. 8) is the middle point of a straight length a of the chosen contour which passes

through two nodal points P, Q of the "net", and let R, S be adjacent nodes, symmetrically arranged with respect to O . Take axes Ox, Oy through S and P , as shown.

The polynomial expression

$$w_0 + Ax + By + Cx^2 + Dy^2 + 2Exy \quad (i)$$

can be substituted for $w(x, y)$ with neglect of terms of order a^3 (at least) as concerns the region $PRQS$; and to this approximation the values of w at S and R (i.e. the points $\pm \sqrt{3}a/2, 0$) are

$$w_S = w_0 + \frac{\sqrt{3}}{2}aA + \frac{3}{4}a^2C, \quad w_R = w_0 - \frac{\sqrt{3}}{2}aA + \frac{3}{4}a^2C,$$

so that $w_S - w_R = \sqrt{3}aA$.

To a corresponding approximation

$$\begin{aligned} \frac{\partial w}{\partial x} &= A + 2Cx + 2Ey, \\ &= A + 2Ey, \text{ for points on the line } QP. \end{aligned} \quad (ii)$$

Therefore to this approximation the contribution of QP to the integral $\int \frac{\partial w}{\partial \nu} ds$ is given by

$$\int_{-\frac{1}{2}a}^{\frac{1}{2}a} \frac{\partial w}{\partial x} ds = aA = (w_S - w_R)/\sqrt{3} \text{ according to (ii).} \quad (31)$$

Provided that the chosen contour consists entirely of straight lines on which nodal points are spaced at equal intervals a , this very simple formula gives the integrals in (30) with an error which is of order a^3 (at least).

Completion of the solution

27. The prediction that

$$I_0 = 2 \frac{A}{L^2} + \int \frac{\partial w_0}{\partial \nu} ds \quad \text{and} \quad I_1 = \int \frac{\partial w_1}{\partial \nu} ds$$

should have the same values for all contours enclosing an internal circle (A being the total included area) was verified by calculations made in accordance with (31) for the three concentric hexagons lettered (a), (b) and (c) in fig. 5. The values obtained were

(a)	(b)	(c)
$I_0 = 0.08865$	0.08916	0.08969
$I_1 = 5.458$	5.452	5.425

and the discrepancies as between contours (a) and (b) can be neglected in an approximate treatment. The results are less reliable for (c), since this contour includes part of the actual boundary.

Taking as mean values

$$I_0 = 0.08917, \quad I_1 = 5.445,$$

we find that

$$k = 0.01638 \quad (32)$$

according to (30). Values of

$$w = w_0 + kw_1,$$

can now be calculated, and these multiplied by L^2 constitute the approximate solution for Ψ . Contours of constant w are shown in the upper part of fig. 7.

IV. AN EXAMPLE IN PLASTICITY: THE TORSION PROBLEM FOR AN OVERSTRAINED SHAFT

28. In the theory of Saint-Venant (Southwell 1936, §389), torsion entails on a cross-section stresses given by

$$X_z = \mu\tau \frac{\partial \Psi}{\partial y}, \quad Y_z = -\mu\tau \frac{\partial \Psi}{\partial x}, \quad (33)$$

where μ denotes the modulus of rigidity and τ the twist per unit length. Ψ is the function introduced in §19: from (33) we see that the shear stress at any point is directed along a line of constant Ψ and has an intensity measured by the gradient of Ψ in a direction perpendicular to that line.

Suppose that the relation between shear stress and strain is linear up to some definite limiting stress f_Y , but that when this stress is attained the strain can increase indefinitely (fig. 9). The assumption accords closely with the phenomenon of "yield" which is exhibited by some materials, and on it Prandtl (1923) has based a theory of torsion beyond the elastic limit. Whereas in the last example Ψ was governed only by boundary conditions combined with (2), now according to our assumption a further restriction is imposed. The gradient of Ψ must not exceed a limiting value $f_Y/\mu\tau$: that is, if w stands for an approximate solution expressed "non-dimensionally" in the manner of §21 ($wL^2 \approx \Psi$), the gradient of w with respect to the non-dimensional coordinates must not exceed a limiting value

$$\lambda = f_Y/\mu L\tau. \quad (34)$$

In regions where it is not governed by this restriction w must satisfy, as before, the finite-difference relation

$$w - \frac{1}{N} \Sigma_{a,N}(w) = \frac{1}{2}a^2, \quad (25 bis)$$

or
$$\frac{1}{N} \left[w \Sigma_{a,N} \left(\frac{1}{x} \right) - \Sigma_{a,N} \left(\frac{w}{x} \right) \right] = \frac{1}{2}a^2 \Sigma_N(x) \quad (29 bis)$$

in relation to a star having unequal arms.

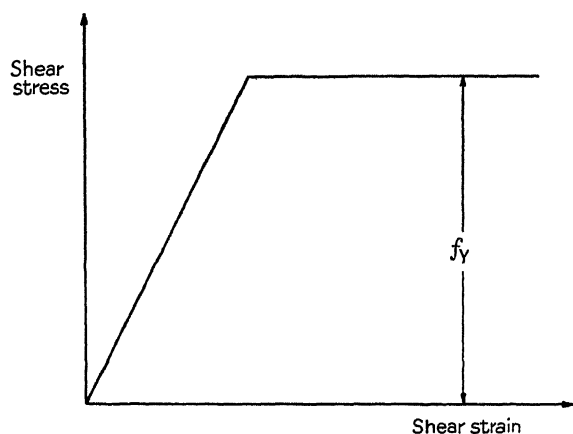


FIG. 9

29. The problem thus presented is intractable by orthodox methods, for the reason that *a priori* we cannot define the boundaries of the regions within which the gradient attains its limiting value. Accordingly Prandtl has proposed an experimental technique which takes advantage of the properties both of soap films and of granular materials with friction. In effect a hole of the requisite shape is "roofed" by a surface of constant slope, and a soap film spread across the hole is deflected by a uniform pressure. In places it bears against the "roof", thus satisfying the overriding condition of § 28: elsewhere it sustains the pressure in virtue of its tension, thus satisfying an equation of the form of (2).*

On the other hand the problem presents no difficulty when treated by relaxation methods. Prandtl's analogy described above shows that the gradient of w will attain its limiting value (if at all) near the boundary: therefore we can postulate values of w which must not be exceeded, and at nodal points where these values are attained the "residual forces" will

* Cf. Nadai (1931), Chap. 19.

not require to be liquidated, because pressure can be sustained by the "roof".

30. No other change in procedure is demanded, and here it will suffice to give results. The section considered was an equilateral triangle, and calculations were made for the following values of λ :

$$(a) 3600\sqrt{3}, \quad (b) 2700\sqrt{3}, \quad (c) 1800\sqrt{3}, \quad (d) 900\sqrt{3}. \quad (35)$$

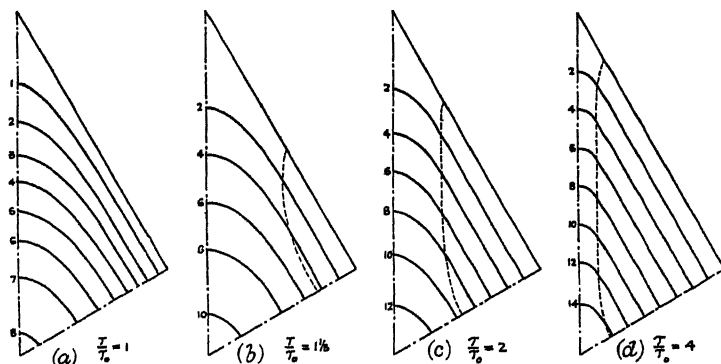


FIG. 10

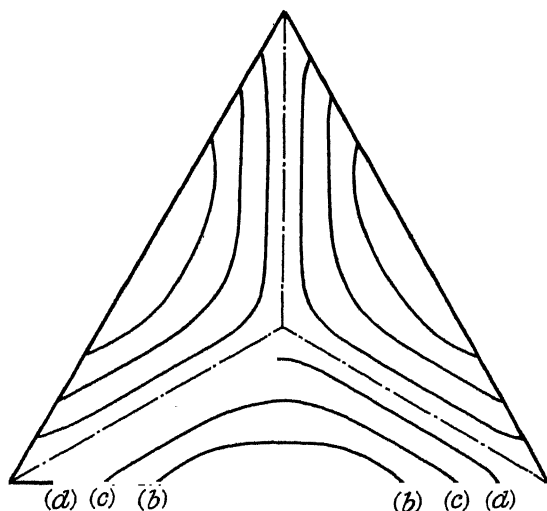


FIG. 11

Fig. 10 exhibits contours of constant w in these four cases, the extent of the overstrained material being shown in each instance by dotted lines. Fig. 11 shows how the overstrained region increases with increasing torque, and

fig. 12 the relation between torque and twist. It is known (Southwell 1936, §§ 339, 390) that the torque

$$T = 2\mu\tau L^2 \iint \Psi dx dy \text{ (when } x \text{ and } y \text{ are "non-dimensional")}$$

$$= 2\mu\tau L^4 \iint w dx dy = \frac{2}{\lambda} L^3 f_Y \iint w dx dy, \text{ according to § 28,} \quad (36)$$

and the surface integral of w is given (with close approximation) by

$$\iint w dx dy = \frac{\sqrt{3}}{2} a^2 \Sigma(w), \quad (37)$$

the summation extending to every nodal point of the triangular net.

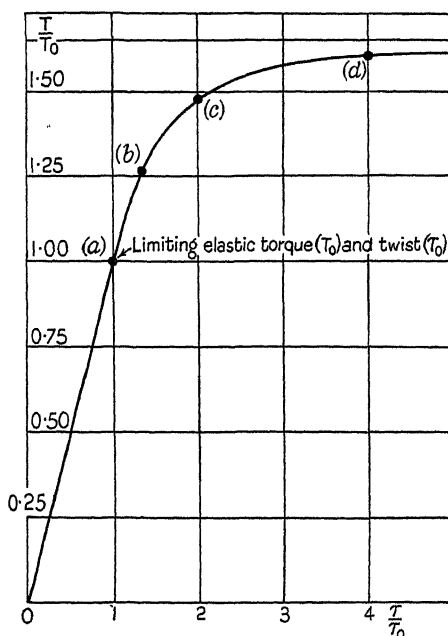


FIG. 12

V. AN EXAMPLE RELATING TO MAGNETIC INDUCTION

31. A magnetic field can be specified by an expression for Ω , the "magnetic potential" (Jeans 1923, § 408). In regions occupied by non-magnetic material the components of magnetic force are

$$\alpha = -\frac{\partial \Omega}{\partial x}, \quad \beta = -\frac{\partial \Omega}{\partial y}, \quad \gamma = -\frac{\partial \Omega}{\partial z}, \quad (38)$$

and in regions occupied by magnetic material of "permeability" μ the components of "magnetic induction" are

$$a = -\mu \frac{\partial \Omega}{\partial x}, \quad b = -\mu \frac{\partial \Omega}{\partial y}, \quad c = -\mu \frac{\partial \Omega}{\partial z}. \quad (39)$$

At every point in the field Ω must be continuous and must satisfy the equation

$$\frac{\partial}{\partial x} \left(\mu \frac{\partial \Omega}{\partial x} \right) + \frac{\partial}{\partial y} \left(\mu \frac{\partial \Omega}{\partial y} \right) + \frac{\partial}{\partial z} \left(\mu \frac{\partial \Omega}{\partial z} \right) = 0, \quad (40)$$

which reduces to Laplace's equation in a region where μ is uniform; and at any point in a surface where the value of μ changes abruptly from μ_1 to μ_2 it must satisfy the boundary condition

$$\mu_1 \frac{\partial \Omega}{\partial \nu_1} + \mu_2 \frac{\partial \Omega}{\partial \nu_2} = 0, \quad (41)$$

ν_1 and ν_2 denoting normals to the surface drawn *into* the two media (Jeans 1923, §§ 432, 466). Exactly similar conditions govern the electrostatic potential in regions which include dielectrics: thus an equation similar to (41) states the law of refraction of lines of magnetic force (Jeans 1923, § 467).

32. Even in two dimensions the calculation of Ω by orthodox methods is difficult, and few solutions exist. Hele-Shaw and Hay devised an experimental technique depending on the approximate analogy of this problem with that of viscous motion between nearly flat and parallel plates (changes in μ being reproduced by changes in the distance between the plates). Their paper (1900) contains many beautiful photographs of results. As a final application of our relaxation methods we now discuss the problem shown in their fig. 29—namely, the effect of an otherwise uniform magnetic field of inserting a triangular prism of iron (permeability 100).

33. Ω being two-dimensional (so that $\partial \Omega / \partial z = 0$), equation (40) permits us to introduce a conjugate function ψ defined by

$$\mu \frac{\partial \Omega}{\partial x} = \frac{\partial \psi}{\partial y}, \quad \mu \frac{\partial \Omega}{\partial y} = -\frac{\partial \psi}{\partial x}. \quad (42)$$

Then on eliminating Ω from (42) we find that ψ (also two-dimensional) is governed by the equation

$$\frac{\partial}{\partial x} \left(\mu^{-1} \frac{\partial \psi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\mu^{-1} \frac{\partial \psi}{\partial y} \right) = 0, \quad (43)$$

so that ψ is plane-harmonic in any region where μ is uniform. The boundary condition (41) will be satisfied provided that no discontinuity occurs in the value of ψ as we pass from the material of permeability μ_1 into the material of permeability μ_2 , but according to (42) the gradient of ψ in a direction normal to the surface of separation must (since Ω is continuous) change its value abruptly in the same ratio as μ : therefore lines of constant ψ will change direction abruptly while they pass into and out of the triangular prism, but ψ will experience no discontinuity.

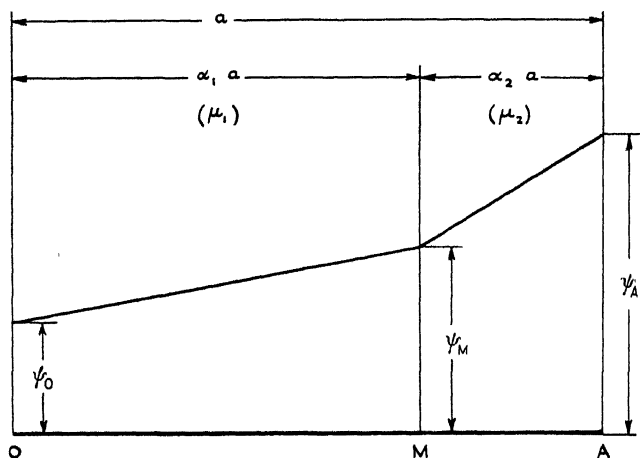


FIG. 13

34. The mechanical analogue of this problem is a transversely loaded membrane in which the tension is proportional to μ^{-1} , and a net will serve as before for an approximation if we assume that every "string" has a tension T proportional to μ^{-1} . Suppose (fig. 13) that a string of length a lies as to a part $(\alpha_1 a)$ of its length in the region of μ_1 and as to the remainder $(\alpha_2 a)$ of its length in the region of μ_2 . If its deflections are ψ_O , ψ_M , ψ_A at O , M and A respectively, the condition for equilibrium at M is

$$\frac{T_2(\psi_A - \psi_M)}{\alpha_2 a} = \frac{T_1(\psi_M - \psi_O)}{\alpha_1 a}, \quad (i)$$

and this serves to determine ψ_M in terms of ψ_O and ψ_A . Then the force exerted by the string on O is given by

$$\frac{T_1(\psi_M - \psi_O)}{\alpha_1 a} = \frac{T_1 T_2(\psi_A - \psi_O)}{a} \bigg/ (T_2 \alpha_1 + T_1 \alpha_2). \quad (ii)$$

The force exerted on A has the same magnitude but opposite sign.

If the whole length OA had been under tension T_1 , the force on O would have been $T_1(\psi_A - \psi_O)/a$. Thus the effect of the altered tension in MA is to alter the effective length of the string from a to

$$a\left(\alpha_1 + \frac{T_1}{T_2}\alpha_2\right) = a\left(\alpha_1 + \frac{\mu_2}{\mu_1}\alpha_2\right), \text{ since } T \propto \mu^{-1},$$

and it follows that in numerical calculation we may employ the formulae (26)–(28) of §§ 23–4, with

$$x = \alpha_1 + \frac{\mu_2}{\mu_1}\alpha_2 \quad (44)$$

in cases where a string crosses from a region where $\mu = \mu_1$ into a region where $\mu = \mu_2$. The argument for this procedure, as for that suggested in § 23, is that it is consistent with the procedure which we have found justification for using in the ordinary case, and that more exact treatment would entail undue complication, considering that at every advance to a finer net we recalculate our errors and (in effect) start afresh.

35. This is the only novelty entailed in the relaxation procedure, and here again it will suffice to give results. Fig. 14, showing calculated contours of constant ψ , is comparable with fig. 29 of the paper by Hele-Shaw and Hay, here reproduced as fig. 15. Except that the experimental method has failed (as always) to reproduce the *sharp* refraction at the boundary of the magnetic material, the contours are practically indistinguishable.

The authors gratefully acknowledge help received from Mr V. Belfield, also from the Secretary and staff of the Aeronautical Research Committee, in the preparation of diagrams suitable for reproduction.

SUMMARY

In this paper relaxation methods are applied with success to four problems involving Laplace's, Poisson's and similar equations in *two variables*, namely:

the torsion problem of Saint-Venant for triangular shafts (this, since the solution is known, serves to illustrate and in some degree to test the methods);

the same problem for a shaft pierced by axial holes (a multiply-connected cross-section);

the torsion problem modified by the imposition of a definite limit to the shear stress (Prandtl's problem);

the problem of magnetic induction in a field containing iron.

In all four problems the methods lead without difficulty to solutions of

sufficient accuracy for practical purposes, and it seems reasonable to conclude that they will be applicable (suitably modified) to other problems which, like the last three, would present great or insuperable difficulties if treated by orthodox methods.

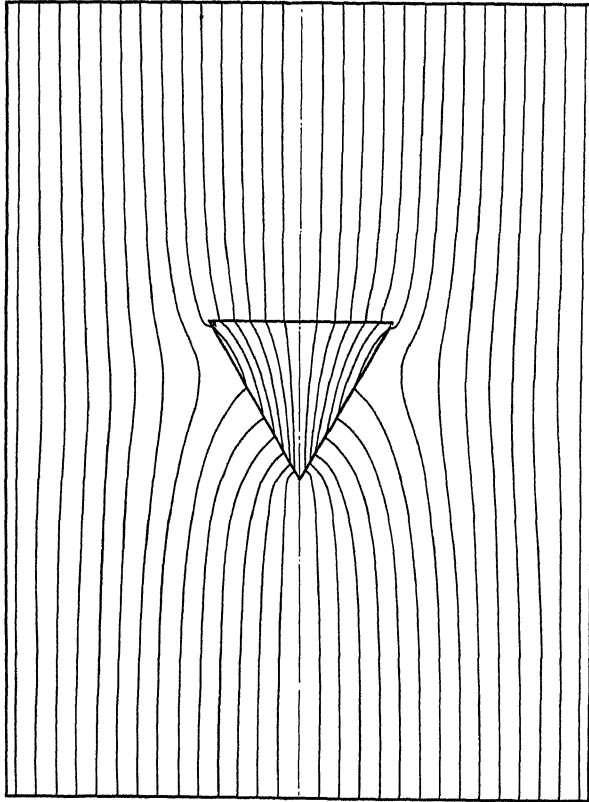


FIG. 14

APPENDIX

On the possibility of formulating a polynomial function of x and y which has specified values at any finite number of specified points

Let lines be drawn parallel with the x -axis to pass through every one of the specified points, and suppose that there are N of these lines, defined by the equations

$$y = y_1, \dots y = y_n, \dots y = y_N. \quad (i)$$

Then, if

$$P(y) = (y - y_1)(y - y_2) \dots (y - y_N), \quad (ii)$$

and if

$$P_n(y) = P(y)/(y - y_n),$$

it is clear that $P_n(y)$ will vanish on all of the lines (i) except the line $(y=y_n)$. Therefore the polynomial function

$$F(x, y) = X_1 P_1(y) + X_2 P_2(y) + \dots + X_n P_n(y) + \dots + X_N P_N(y) \quad (\text{iii})$$

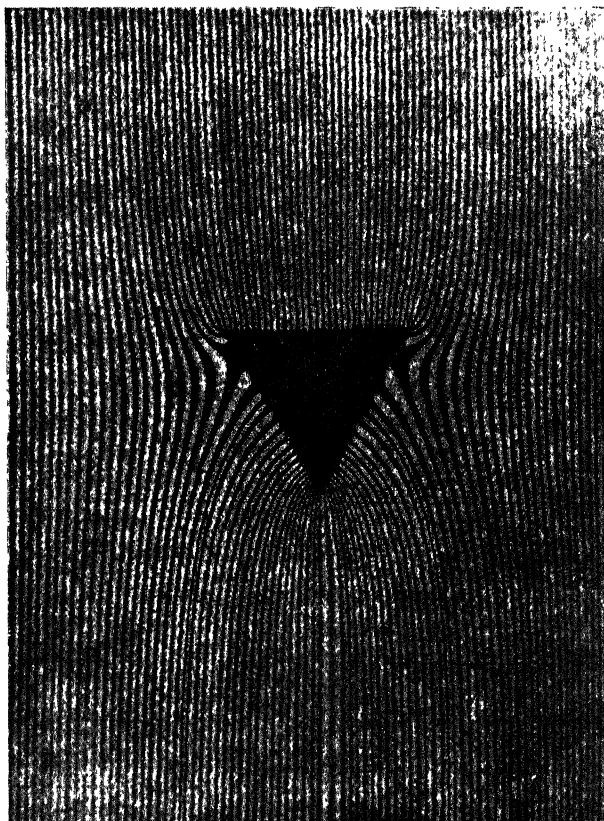


FIG. 15

(in which X_1, X_2, \dots, X_N are functions of x as yet unspecified) will be given on the line $y = y_n$ by

$$F(x, y)_n = X_n P_n(y_n). \quad (\text{iv})$$

Suppose that at those specified points which lie on the line $(y = y_n)$ x has the values a_n, b_n, \dots, k_n , and that the specified values of the polynomial function at these points are $Q_{a,n}, Q_{b,n}, \dots, Q_{k,n}$. Then if

$$\Pi_n(x) = (x - a_n)(x - b_n) \dots (x - k_n), \quad \left. \vphantom{\Pi_n(x)} \right\} \quad (\text{v})$$

and if

$$\Pi_{a,n}(x) = \Pi_n(x - a_n),$$

it is clear that $\Pi_{a,n}(x)$ vanishes at all of the points $(x = b_n), \dots (x = k_n)$, but at the point $(x = a_n)$ has a finite value

$$\Pi_{a,n}(a_n) = (a_n - b_n) \dots (a_n - k_n). \quad (\text{vi})$$

Therefore the polynomial function of x given by

$$X_n = \frac{1}{P_n(y_n)} \left[Q_{a,n} \Pi_{a,n}(x) / \Pi_{a,n}(a_n) + \dots + Q_{k,n} \Pi_{k,n}(x) / \Pi_{k,n}(k_n) \right], \quad (\text{vii})$$

will assume the values

$$\frac{1}{P_n(y_n)} [Q_{a,n}, \dots, Q_{k,n}] \quad (\text{viii})$$

at the points a_n, \dots, k_n on the line $y = y_n$; and it follows from (iv) that if we substitute the value (vii) for X_n in (iii), then $F(x, y_n)$ will assume the specified values at the specified points.

Deriving expressions of the type of (vii) for all of X_1, X_2, \dots, X_N in (iii), we shall have an expression for $F(x, y)$, polynomial both in x and y , which takes the specified value at every one of the specified points.

REFERENCES

- Baird, L. and Berry, A. 1919 *Proc. Roy. Soc. A*, **95**, 457.
 Baird, L., Cave, B. M. and Lang, E. D. 1923 *Philos. Trans. A*, **223**, 383.
 Baird, L. and Pippard, A. J. S. 1922 *Proc. Instn Civ. Engrs*, **214**, 291.
 Black, A. N. and Southwell, R. V. 1938 *Proc. Roy. Soc. A*, **164**, 447.
 Bradfield, K. N. E., Hooker, S. G. and Southwell, R. V. 1937 *Proc. Roy. Soc. A*, **159**, 315.
 Bradfield, K. N. E. and Southwell, R. V. 1937 *Proc. Roy. Soc. A*, **161**, 155.
 Frazer, R. A. 1937 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1799.
 Hele-Shaw, H. S. and Hay, A. 1900 *Philos. Trans. A*, **195**, 303.
 Jeans, J. H. 1923 "The Mathematical Theory of Electricity and Magnetism", 4th ed. Camb. Univ. Press.
 Love, A. E. H. 1927 "The Mathematical Theory of Elasticity", 4th ed. Camb. Univ. Press.
 Nadai, A. 1931 "Plasticity". McGraw Hill Co.
 Prandtl, L. 1903 *Phys. Z.* **4**, 758.
 — 1923 *Z. angew. Math. Mech.* **3**, No. 6, 442.
 Relf, E. 1924 *Phil. Mag.* (6), **48**, 535; also *Rep. Memor. Aero. Res. Comm., Lond.*, no. 905.
 Southwell, R. V. 1936 "Introduction to the Theory of Elasticity". Oxford Univ. Press.
 Taylor, G. I. and Griffith, A. A. 1917 *Proc. Instn Mech. Engrs, Lond.*, p. 755.
 Thom, A. 1928 *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1194.
 Whittaker, E. T. and Robinson, G. 1926 "The Calculus of Observations", 2nd ed. London: Blackie and Son.

Oscillatory motion of a fluid along a circular tube

By D. G. CHRISTOPHERSON, A. GEMANT, A. H. A. HOGG
AND R. V. SOUTHWELL, F.R.S.

(Received 26 July 1938)

INTRODUCTORY

1. This paper deals with oscillatory motions of a viscous incompressible fluid in a uniform circular tube. The oscillations may either be free or forced by a pulsating pressure, control coming in both instances from the "gravity head" which results from the passage of fluid. They are closely analogous with mechanical and electrical oscillations, but the theory is complicated by the circumstance that the fluid does not oscillate as a rigid body, its velocity ranging between a zero value at the wall and a maximum value at the axis of the tube.

Our interest in the problem originated in a notion that oscillatory motion, either "free" or "forced", might be made the basis of a new method of viscometry. Of existing methods, all which aim at high accuracy appear (following the pioneer method of Poiseuille 1840-1) to take advantage of the known solution for steady flow along a straight and uniform tube. Commercial viscometers employ the same principle, but the tube (for convenience) is made relatively short, and the motion is not steady, since the "head" of the fluid is changing throughout the tests.* On this account their measurements are comparative rather than absolute; for since the motion is not calculable by theory they must be calibrated with the aid of fluids whose viscosity is known. They have the advantage (compared with Poiseuille's method) of not requiring a large amount of the fluid under test; but it is difficult to be certain of the exact temperature of the oil while it is passing through the tubular orifice, and this is a serious disadvantage in view of the rapidity with which the viscosity of oils decreases with increase of temperature.† Our notion was that a method involving pulsating instead of steady flow on the one hand would call for only a small amount of the fluid under test and on the other would have some claim to be regarded as an absolute method. This is not the place for a complete account of the resulting "oscillation viscometer", but such features as have a bearing

* The "ixometre" of Barbey is an exception (cf. Barr 1931, chap. iv).

† Barr 1931, chap. i, § 6.

on the fundamental problem of this paper receive brief notice in Part II; Part I deals with theory and Part III with experimental results.

We acknowledge with gratitude help received in the preparation of diagrams for this paper from the Secretary of the Aeronautical Research Committee and his staff, and from our colleague Mr V. Belfield.

I. THEORY

2. Simplified with a view to theoretical treatment our problem may be stated as follows (fig. 1): The fluid (here treated as incompressible and characterized by density ρ and viscosity μ) executes rectilinear oscillations in the straight tube AB , starting from and discharging to reservoirs A and B . Discharge into A will cause the free surface in that reservoir to rise and the free surface in B to fall: it is resisted by a gravitational pressure-difference $2g\rho x$, $2x$ being the height by which the level of A exceeds the level of B . We assume that this pressure difference is transmitted without change to the terminal sections of the tube, and has there the same value at all radii: that is to say, we calculate the motion of the fluid column on the assumption that the net pressure difference tending to maintain the displacement is

$$p - 2g\rho x$$

at all radii, p denoting the externally applied pressure in the direction of the excursion x .

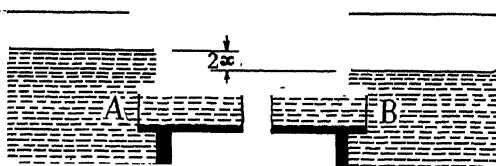


FIG. 1

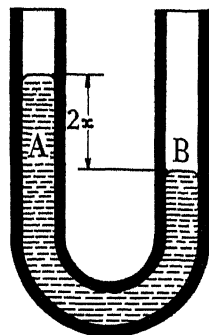


FIG. 2

We assume that v , the velocity of the fluid along the tube, is a function of time (t) and of the radial distance (r) but is independent of angular position (θ) in the cross-section. We assume further that the velocity has no radial component, so that v has the same distribution in all cross-sections. Near the ends, of course, radial velocities must exist on account of the motion

in the reservoirs; but they can have little *dynamic* effect, and we may expect that in a reasonably long column this last assumption will be close to the truth.

3. We form the equation of motion for an annular cylinder of fluid having radii r and $r + \delta r$, with length $(2h)$ equal to that of the fluid column. Its mass is $\rho \times 2h \times 2\pi r \delta r$, and the frictional force is $2h\delta r \frac{\partial}{\partial r} \left(2\pi r \mu \frac{\partial v}{\partial r} \right)$, its direction being that of v . The pressure difference exerts a force $2\pi r \delta r \times (p - 2g\rho x)$ in the same direction: therefore the equation of motion is

$$\ddot{v} = \frac{\mu}{\rho} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v}{\partial r} \right) + \left(\frac{p}{2\rho h} - g \frac{x}{h} \right). \quad (1)$$

Now let A stand for the area of the free surface in each reservoir. (We assume that the reservoirs are cylindrical, so that A is independent of the displacement.) Then evidently

$$\begin{aligned} A \frac{dx}{dt} &= \text{total flow through the straight tube} \\ &= \int_0^a v 2\pi r dr, \end{aligned} \quad (2)$$

and so, after differentiating (1) with respect to time, we can eliminate x with the aid of (2). The result is

$$\ddot{v} - \frac{\mu}{\rho} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \dot{v}}{\partial r} \right) + \frac{2\pi g}{hA} \int_0^a r v dr = \frac{\dot{p}}{2\rho h}, \quad (3)$$

dots denoting differentiations with respect to t .*

$$\left. \begin{aligned} 4. \text{ Assuming now that } & p = P e^{\text{Int}}, \\ & x = X e^{\text{Int}}, \\ & v = V e^{\text{Int}}, \end{aligned} \right\} \quad (4)$$

* Increase of p by a constant (steady) pressure p_0 entails a constant addition x_0 to x , where

$$p_0 = 2g\rho x_0;$$

v then vanishes according to (2), and (3) is satisfied. Hence, if p fluctuates about a non-zero mean, the free surfaces will fluctuate about mean positions removed by distances x_0 from their positions of equilibrium when both free surfaces are exposed to the same pressure.

where P and X are constant and V a function of r only, we substitute in (2) and (3) to obtain

$$\left. \begin{aligned} \frac{d^2 V}{dr^2} + \frac{1}{r} \frac{dV}{dr} + k^2 V &= Q - \frac{P}{2h\mu}, \\ \text{where} \quad k^2 &= -in\rho/\mu, \\ Q &= \frac{2\pi g}{hA} \frac{k^2}{n^2} \int_0^a rVdr. \end{aligned} \right\} \quad (5)$$

Accordingly we have

$$\begin{aligned} \frac{hn^2}{g'} Q &= \frac{2k^2}{a^2} \int_0^a rVdr = \frac{2}{a^2} \int_0^a \left\{ r \left(Q - \frac{P}{2h\mu} \right) - \frac{d}{dr} \left(r \frac{dV}{dr} \right) \right\} dr, \\ &= Q - \frac{P}{2h\mu} - \frac{2}{a^2} \left[r \frac{dV}{dr} \right]_0^a, \end{aligned} \quad (6)$$

$$\text{where} \quad g' = \frac{\pi a^2}{A} g. \quad (7)$$

Equation (7) gives the ratio in which gravity is reduced (effectively) by reason of the fact that our straight tube (fig. 1) discharges at each end into a reservoir of greater cross-sectional area A . On certain assumptions the theory now being formulated may be regarded as applying to oscillations in a U-tube of uniform circular bore (fig. 2): then $A = \pi a^2$, the cross-sectional area of the tube, and so

$$g' = g$$

according to (7). With g' defined by (7) the relation (6) applies generally.

The assumptions in question are

(a) that the curvature of the U-tube may be neglected as concerns its effect on the motion of the fluid; i.e. that the fluid velocity (as in the straight tube) is directed everywhere along the central line of the tube and is a function only of r , the distance from the central line;

(b) that surface tension operates to keep both free surfaces flat; i.e. that the pressure-difference may be taken as independent of position in the cross-section.

5. The solution of (5) which satisfies the condition of no "slip" at the boundary is

$$V = \frac{1}{k^2} \left(Q - \frac{P}{2h\mu} \right) \left[1 - \frac{J_0(kr)}{J_0(ka)} \right]. \quad (8)$$

$$\text{Then} \quad \frac{dV}{dr} = -\frac{1}{k} \left(Q - \frac{P}{2h\mu} \right) \frac{J'_0(kr)}{J_0(ka)} = \frac{1}{k} \left(Q - \frac{P}{2h\mu} \right) \frac{J_1(kr)}{J_0(ka)},$$

and we have from (6)

$$\left[\frac{g'}{hn^2} - \left\{ 1 - \frac{2}{ka} \frac{J_1(ka)}{J_0(ka)} \right\}^{-1} \right] Q = \frac{Pg'}{2h^2\mu n^2}.$$

But from (2), (4) and (5) we see that

$$X = \frac{h\mu}{g\rho} Q. \tag{9}$$

Therefore we have, finally,

$$\left[\frac{g'}{hn^2} - \left\{ 1 - \frac{2}{ka} \frac{J_1(ka)}{J_0(ka)} \right\}^{-1} \right] X = \frac{g'}{g} \frac{P}{2\rho hn^2} \tag{10}$$

as an equation relating the applied (periodic) pressure and the resulting (periodic) displacement of the free surfaces. In (10) both P and X (in general) are complex quantities.

Free oscillations

6. When P is zero we have an equation determining the period and decay factor for free oscillations. It may be written in the form

$$\left. \begin{aligned} \frac{hn^2}{g'} &= 1 - \frac{2J_1(ka)}{kaJ_0(ka)}, \\ \frac{h}{g'} \left(\frac{\mu z^2}{\rho a^2} \right)^2 &= 1 - \frac{2J_1(iz\sqrt{i})}{iz\sqrt{i}J_0(iz\sqrt{i})}, \end{aligned} \right\} \tag{11}$$

or

where, by the second of (5)

$$z^4 = -k^4 a^4 = \frac{n^2 \rho^2 a^4}{\mu^2}. \tag{12}$$

We know (since free oscillations must decay) that in the general case n and therefore z will be complex: therefore in the absence of tables relating to Bessel functions of complex argument it would seem that a complete solution of (11) is out of the question.

We can, however, solve (11) if we assume either that z is very large or very small. Evidently z can have a real value

$$\bar{z} = a \sqrt[4]{\frac{g'\rho^2}{h\mu^2}} \tag{13}$$

provided that the term involving Bessel functions is negligible; and for real values of z we can show (cf. § 8 and Table I) that

$$\frac{2J_1(iz\sqrt{i})}{iz\sqrt{i}J_0(iz\sqrt{i})} = \{A(z) + iB(z)\}^{-1},$$

where $A(z)$ and $B(z)$ are functions of z increasing steadily from 1 and 0 respectively to a limiting value $z/\sqrt{8}$ (cf. § 8) as z increases from 0 to ∞ . Accordingly z can have a real value \bar{z} provided that this is very large, and from (13) we see that \bar{z} will be very large provided that μ/ρ is very small. So, when the viscosity is negligible, the equation of free oscillations tends to the limiting form

$$n = \sqrt{g'/h}. \quad (14)$$

We can now deduce an approximate solution* for the case of finite but very small viscosity by substituting for the right-hand side of (11) its first approximation when z is large, viz.

$$1 - \{(1+i)\bar{z}/\sqrt{8}\}^{-1} = 1 - \frac{(1-i)\sqrt{2}}{\bar{z}}.$$

Then, in place of (14), we have as a second approximation

$$n = \sqrt{\frac{g'}{h}} \left\{ 1 - \frac{(1-i)\sqrt{2}}{\bar{z}} \right\}^{\frac{1}{2}} = \sqrt{\frac{g'}{h}} \left\{ 1 - \frac{1}{\bar{z}\sqrt{2}} + \frac{i}{\bar{z}\sqrt{2}} \right\}, \quad (15)$$

when the second and higher powers of $1/\bar{z}$ are neglected. This indicates that on account of friction the period of a complete oscillation will be lengthened from

$$T_0 = 2\pi\sqrt{\frac{h}{g'}} \quad \text{to} \quad T = T_0 \left(1 - \frac{1}{\bar{z}\sqrt{2}} \right)^{-1} = T_0 \left(1 + \frac{1}{\bar{z}\sqrt{2}} \right)$$

to the same approximation; and that the ratio of one excursion to the next (on the opposite side) will be $e^{\lambda T/2}$, where by (4) and (15)

$$\lambda = \frac{1}{\bar{z}} \sqrt{\frac{g'}{2h}}.$$

The logarithm of the ratio (i.e. the "logarithmic decrement" of the oscillation) is

$$\begin{aligned} \delta &= \frac{1}{2}\lambda T = \frac{T_0}{2\bar{z}} \sqrt{\frac{g'}{2h}}, \text{ to the first order in } 1/\bar{z}, \\ &= \frac{\pi}{\bar{z}\sqrt{2}} = \frac{\pi}{a} \sqrt{\frac{h\mu^2}{4g'\rho^2}}, \text{ by (13).} \end{aligned} \quad (16)$$

These formulae should be applicable to mercury, for which $\mu/\rho = 0.0012$, approximately, in c.g.s. units. For unless the bore of the tube is extremely small, friction may be expected to have very little influence even when allowance is made for the circumstance (disregarded in this theory) that

* (Added 13 Aug. 1938.) Somewhat similar analysis has been applied by Fage (1938) to the problem of a viscous fluid oscillating axially in an *annular* pipe.

rectilinear motion may become unstable and give place (locally at least) to turbulent flow.

7. At the other end of the scale, when z is small, we may use the series expressions

$$J_1(x) = \frac{x}{2} \left\{ 1 - \frac{(\frac{1}{2}x)^2}{1 \cdot 2} + \frac{(\frac{1}{2}x)^4}{1 \cdot 2 \cdot 2 \cdot 3} - \dots \right\},$$

$$J_0(x) = 1 - \frac{(\frac{1}{2}x)^2}{1!^2} + \frac{(\frac{1}{2}x)^4}{2!^2} - \dots$$

Then, if we neglect powers of z higher than the fourth in comparison with unity, (11) reduces to

$$1 - \frac{8}{i} \frac{\hbar}{g'} \left(\frac{\mu}{\rho a^2} \right)^2 z^2 \left(1 + \frac{iz^2}{6} \right) = 0,$$

or to

$$-n^2 + 2\lambda i n + \alpha^2 = 0,$$

where

$$\lambda = \frac{3\mu}{\rho a^2}, \quad \alpha^2 = \frac{3g'}{4\hbar}.$$

According to this equation, the approximate solution for n is

$$i\lambda \pm \sqrt{(\alpha^2 - \lambda^2)},$$

and hence, according to (12),

$$z^2 = 3\{i \pm \sqrt{(\alpha^2/\lambda^2 - 1)}\}.$$

It follows that

$$|z^2| > 3 \quad \text{if} \quad \alpha^2 > \lambda^2,$$

as must be the fact in order that oscillations (as distinct from a subsidence) may occur: therefore, when the conditions are those in which we are interested, the magnitude of z will be such as to throw doubt on the approximation which results from the assumption that $|z|$ is small.

It thus appears that little can be done in theory to predict the behaviour of free oscillations which are controlled to an appreciable extent by friction. For this reason (among others; cf. § 12), greater practical importance attaches to the theory of forced oscillations.

Forced oscillations

8. When the oscillations are forced by a periodic pressure, n is a real constant and accordingly k^2 , as defined by the second of (5), is a pure imaginary. Introducing the quantity z as before (§ 6), we may write (10) in the form

$$\left[\frac{g'}{\hbar n^2} - \left\{ 1 - \frac{2J_1(iz\sqrt{i})}{iz\sqrt{i}J_0(iz\sqrt{i})} \right\}^{-1} \right] X = \frac{g'}{g} \frac{P}{2\rho\hbar n^2}, \quad (17)$$

where z is real. The quantity in twisted brackets may be tabulated with the aid of tables of the ber and bei functions first introduced by Lord Kelvin (1890; cf. Watson 1923 and Jahnke and Emde 1909). Writing

$$J_0(iz\sqrt{i}) = \text{ber}(z) + i \text{bei}(z),$$

we have
$$-i\sqrt{i}J_1(iz\sqrt{i}) = \frac{d}{dz}J_0(iz\sqrt{i}) = \text{ber}'(z) + i \text{bei}'(z),$$

so that
$$\frac{iz\sqrt{i}J_0(iz\sqrt{i})}{2J_1(iz\sqrt{i})} = \frac{iz}{2} \frac{\text{ber}(z) + i \text{bei}(z)}{\text{ber}'(z) + i \text{bei}'(z)},$$
$$= \frac{z}{2} \frac{W(z) + iZ(z)}{V(z)}, \quad (18)$$

where V , W , Z stand for functions defined as under (Russell 1914):

$$\left. \begin{aligned} V(x) &= \text{ber}'^2 x + \text{bei}'^2 x, \\ Z(x) &= \text{ber } x \text{ber}' x + \text{bei } x \text{bei}' x, \\ W(x) &= \text{ber } x \text{bei}' x - \text{bei } x \text{ber}' x. \end{aligned} \right\} \quad (19)$$

Accordingly we have

$$\left\{ 1 - \frac{2J_1(iz\sqrt{i})}{iz\sqrt{i}J_0(iz\sqrt{i})} \right\}^{-1} = [1 - \{A(z) + iB(z)\}^{-1}]^{-1} \\ = C(z) - iD(z),$$

where A , B , C and D are functions defined by

$$\left. \begin{aligned} A(x) &= \frac{x}{2} \frac{W(x)}{V(x)}, & C(x) &= 1 + \frac{A(x) - 1}{\{A(x) - 1\}^2 + B(x)^2}, \\ B(x) &= \frac{x}{2} \frac{Z(x)}{V(x)}, & D(x) &= \frac{B(x)}{\{A(x) - 1\}^2 + B(x)^2}, \end{aligned} \right\} \quad (20)$$

and equation (17) may thus be written in the form

$$\left[\frac{g'}{\hbar n^2} - C(z) + iD(z) \right] X = \frac{g'}{g} \frac{P}{2\rho\hbar n^2}. \quad (21)$$

The factor enclosed within square brackets is a complex quantity defining both the amplitude and phase of the fluid oscillation. The functions $A(z)$ and $B(z)$ are the functions denoted by w'/w_0 and $L'_i\omega/w_0$, respectively, in the "Funktionentafeln" of Jahnke and Emde, where the independent variable is denoted by q .^{*} As z tends to ∞ (fluid of very small viscosity),

^{*} Section XIII, § 7, Table XXI and Fig. 42 of the first edition (Berlin, 1909), pp. 144-7.

$A(z)$ and $B(z)$ both tend to the value $z/\sqrt{8}$, so that $C(z)$ tends to 1 and $D(z)$ to zero: then we have as the limiting form of (21)

$$\left(1 - \frac{h}{g'} n^2\right) X = \frac{P}{2g\rho}, \quad (22)$$

—an equation which is evidently correct when $n = 0$, and which indicates that “resonance” will occur when the period of a pulsation is that of a simple pendulum of length hg/g' . This result was to be expected.

9. At the other end of the scale (fluids of high viscosity, tested in tubes of small bore) we have, proceeding as in § 7,

$$\begin{aligned} C(z) - iD(z) &= \left\{1 - \frac{2J_1(iz\sqrt{i})}{iz\sqrt{i}J_0(iz\sqrt{i})}\right\}^{-1} \\ &= \frac{4}{3} \left(1 - \frac{z^4}{16 \times 360}\right) - \frac{8i}{z^2} \left(1 + \frac{z^4}{16 \times 72}\right) \end{aligned} \quad (23)$$

if we neglect terms of order z^8 in comparison with unity: so when z is small the imaginary part predominates in the quantity $C(z) - iD(z)$. With neglect of positive powers of z equation (21) becomes

$$\left[1 - \frac{4hn^2}{3g'} + 8i\frac{h\mu n}{g'\rho a^2}\right] X = \frac{P}{2g\rho}, \quad (24)$$

when we substitute for z^2 from (12). This shows that at very low frequencies the response (as we should expect) is $P/2g\rho$; but unless n is very small the viscous term will predominate in the square brackets, and then we shall have approximately

$$\frac{g}{g'} \frac{|X|}{|P|} = \frac{1}{16} \frac{a^2}{\mu hn}, \quad (25)$$

i.e. a response inversely proportional to the frequency of the pulsating pressure.

Equations (22) and (25) are the theoretical basis of our “forced oscillation method” of viscometry.

10. For values of z between 2.5 and 20, $C(z)$ and $D(z)$ may be calculated, by (20), from values of $A(z)$ and $B(z)$ taken (cf. § 8) from Table XXI of the “Funktionentafeln”. For smaller values of z the approximate formula (23) can be used. Values obtained in this way are given below, and the trend of the functions in fig. 3.

TABLE I. VALUES OF THE FUNCTIONS $C(z)$ AND $D(z)$

z	$A(z) - 1$	$B(z)$	$C(z)$	$D(z)$
0	0	0	1.3333	∞
0.5	0.0000	0.0313	1.3333	31.949
1.0	0.0001	0.1247	1.3331	8.0192
1.5	0.0258	0.2777	1.3322	3.5702
2.0	0.0805	0.4806	1.3296	2.0239
2.5	0.1747	0.7136	{ 1.3243 1.3237 }	1.3221
3.0	0.3180	0.9508	1.3164	0.9459
3.5	0.4920	1.172	1.3045	0.7254
4.0	0.6778	1.373	1.2898	0.5870
4.5	0.8628	1.558	1.2720	0.4912
5.0	1.0430	1.737	1.2541	0.4231
5.5	1.2190	1.916	1.2364	0.3715
6	1.3937	2.093	1.2204	0.3310
8	2.0956	2.814	1.1702	0.2286
10	2.7940	3.487	1.1399	0.1746 _s
15	4.5732	5.316	1.0931	0.1081
20	6.3250	7.082	1.0701 _s	0.0785 _s

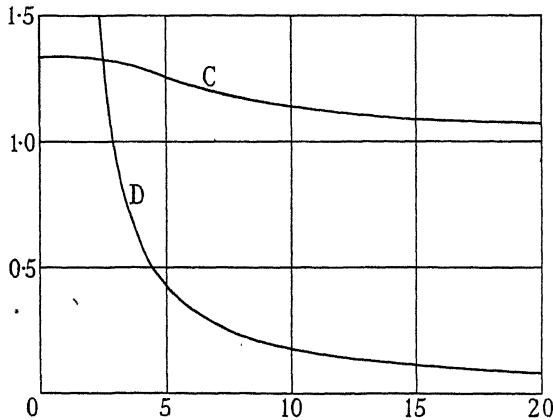


FIG. 3

Using this table we can construct a vector diagram on the lines of fig. 4 in which AB represents, for an origin at Q , the end of a variable vector $QP = \{-C(z) + iD(z)\}$. If we take O to the left of Q by an amount representing g'/hn^2 , then OP will represent, for the value of z corresponding with any point P on AB , the quantity included within the square brackets in (21). The displacement will "lag" behind the pressure by a phase-angle POQ , and its amplitude will be given by

$$|X| = \frac{|P|}{2\rho hn^2} \frac{g'}{|OP|g} = \frac{|OQ|}{|OP|} \times \frac{|P|}{2g\rho}, \quad (26)$$

in which $|P|$ denotes the amplitude of the pressure p according to (4). We observe that $|P|/2g\rho$ would be the response of the fluid to a *maintained* pressure $|P|$.

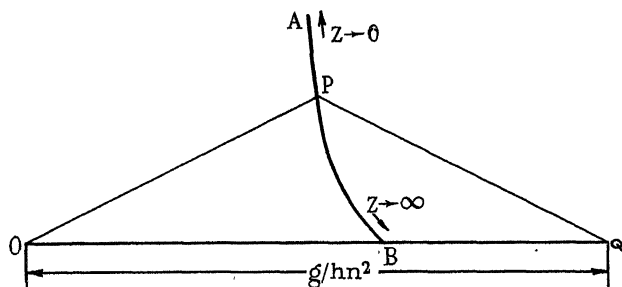


FIG. 4

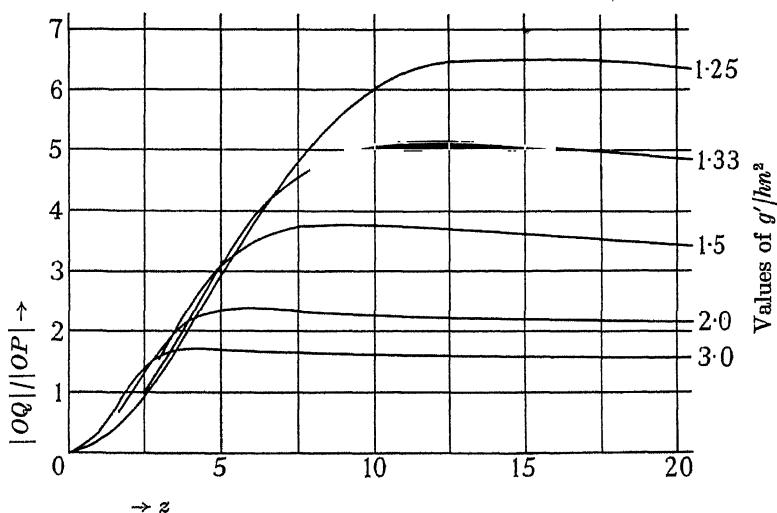


FIG. 5

Fig. 5 has been constructed from a large-scale plot of fig. 4 to show, for the range $0 < z < 20$, the variation with z of $|OQ|/|OP|$ when g'/hn^2 (represented by $|OQ|$) has the values 1.25, 1.33, 1.5, 2.0 and 3.0. It indicates that large values of g'/hn^2 will give a response which is insensitive to variations of viscosity over practically the whole range, but is small: on the other hand it suggests that we may make g'/hn^2 as small as 1.5 or even 1.25, and thus obtain the same insensitivity in a response which is greater, provided we can arrange that z is not less than about 12. The same conclusion can be drawn directly from equation (21); for when z exceeds 15 or 20 the quantities

$\{C(z) - 1\}$ and $D(z)$ are both represented closely by $\sqrt{2/z}$, and hence, approximately, that equation reduces to

$$\left(1 - \frac{hn^2}{g'}\right) X = \frac{P}{2g\rho}, \quad (22 \text{ bis})$$

which makes X independent of changes in the viscosity.

II. PRACTICAL CONSIDERATIONS

11. To summarize this theoretical investigation, we have assumed that x , which measures the displacement of the free surface, and p , the applied pressure, both involve the time as an exponential time-factor e^{int} in which n may be either real or imaginary. On this assumption we have obtained the relation (10) between the applied pressure and the response thereto. When $P = 0$ we have the equation (11) governing free oscillations. This can be solved approximately if \hat{z} is small (i.e. when the viscosity of the fluid is large), but (§ 7) the approximation is not reliable except in circumstances where a subsidence, as distinct from damped oscillations, is to be expected. It can also be solved approximately when z is large, and for fluids of negligible viscosity it reduces to (14), indicating (as we should expect) a simple harmonic displacement of period $2\pi\sqrt{h/g'}$.

When the friction is small there is a possibility (not contemplated in the theoretical investigation) that rectilinear flow may be replaced—at all events in part of the tube—by a turbulent regime. This is not of much importance when, as in the case of mercury oscillating in a tube of fairly large bore, the motion will in any event be almost unaffected by viscosity; but the possibility must be kept in mind when the formulae for *forced* oscillations are applied to viscometry.

Reynolds' number R , the criterion of transition from laminar to turbulent flow, will be proportional to the product of a representative length and velocity divided by the kinematic viscosity $\nu (= \mu/\rho)$ of the fluid. If we take a and $\frac{g}{g'}n|X|$ as the representative length and velocity,* then

$$R = \frac{g}{g'}\rho an|X|/\mu = \frac{g}{g'}z^2|X|/a, \quad \text{by (12),}$$

and it would thus appear that turbulence can be avoided either by using tubes of small bore or by restricting the amplitude of the excursion. It does

* $\frac{g}{g'}|X|$ is a measure of the fluid displacement *in the tube*.

not seem possible to predict the critical value of Reynolds' number, and no clear indication is afforded by published experiments which, almost without exception, relate to the stability of *steady* flow down *long* tubes;* but such evidence as is available suggests that it would not be safe to work with values of z in excess of 1. This of course is a question to be settled by experiment, as also is the question of end effects (which our theory disregards).

12. In thinking of oscillatory motions as providing the basis of a possible method of viscometry, we contemplated that the oscillations might be either "free" or "forced". In the "free oscillation method"

(fig. 6) two U-tubes are placed in series with a small air space separating the two fluid columns, one tube being filled with mercury, the other with the fluid under test. Tested separately, the mercury, on account of its high density and low viscosity, would oscillate freely for a long while, but the oil would soon come to rest: placed in series, the two fluids oscillate

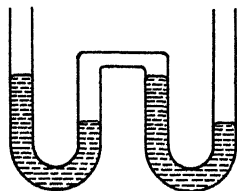


FIG. 6

with a decay factor which can be regulated by altering the length of the mercury column.† The measured decay factor would be interpreted by means of the formulae of §§ 6–7, which are easily extended to the case of two U-tubes in series. (This is evident from the form of (21), in which the pressure difference P and displacement X can be interpreted electrically as potential difference and current: then the quantity in square brackets (appropriately multiplied) is the analogue of "impedance", and we can deal with U-tubes in series or in parallel exactly as in the theory of electric networks.)

But in practice the "free oscillation method" is restricted: for unless viscosity exerts a predominating influence results are difficult to interpret; and if it does, then the oscillations decay rapidly and *measurement* is difficult. A few preliminary experiments led us to discard this first alternative.

13. In the "forced oscillation method" (believed to be new) the two U-tubes are placed "in parallel" instead of "in series", and the same pulsating pressure is applied to both. Fig. 7 shows (diagrammatically) the general arrangement. Temperature control is easy, and very little fluid is required. Because the motion is maintained, observations can be repeated

* Cf. the second footnote to § 23.

† Menneret (1911) employed the method of free oscillations for the measurement of viscosity, but does not appear to have used our device of a mercury tube in series (cf. Barr 1931, chap. x).

indefinitely; and because in this instance an approximate theory can be constructed, they may be analysed either on an absolute or on a comparative basis. Measurement of pressure is not necessary, since the viscosity of the fluid under test can be determined from the ratio of the two amplitudes (exactly as in electricity two impedances can be compared by measuring the currents which flow through them when arranged in parallel). The response of the mercury tube can be calculated from (22), and the response of the fluid under test from (25): then we have two responses related with the same quantity $|P|$, and eliminating $|P|$ we have a relation whereby results can be interpreted.

This method is the basis of our "oscillation viscometer". Detailed description is reserved for a subsequent paper.*

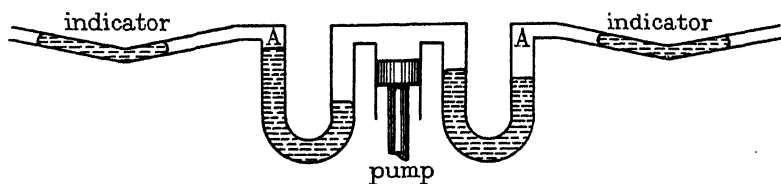


FIG. 7

III. EXPERIMENTAL

14. We now describe experiments which have been made to check the theory given in Part I.

Experiments using free oscillations

The aim of these experiments was to test the conclusions of § 6 as applied to free oscillations of mercury in a bent glass U-tube, and in particular to study the effects of surface tension. It was also thought that they might throw some light on the question (§ 11) whether our formula (22) for the response of mercury to a pulsating pressure is likely to be invalidated by the onset of turbulence at high values of Reynolds' number. The work resulted in a satisfactory technique for the observation of damped free oscillations, and although the experiments were few it seems worth while to put this on record.

Timing by stop-watch was found to be insufficiently exact, on account of the small number of oscillations large enough to be measured accurately; but by using a small cinematograph to photograph the apparatus together with a "clock" (the latter a synchronous motor making approximately

* The instrument is covered by British Patent No. 434994 (1934).

3 rev./sec.) we were able to construct diagrams of which fig. 9 is typical, showing the variation of the excursion (x) on a known scale of time. For this purpose the cinematograph record (exemplified by fig. 8) was examined in a microscope fitted with a hair-line which could be traversed by a micrometer screw. The "clock" readings were recorded in tables, along with the instantaneous positions (as read on the micrometer) of the two ends of the fluid column. For determining periods and damping factors it was not necessary to translate the micrometer readings into absolute values of the excursion.

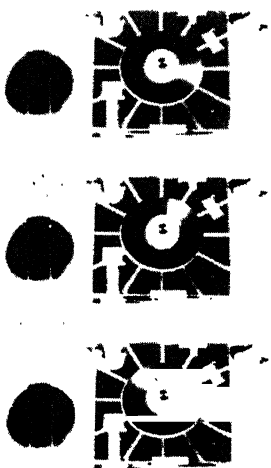


FIG. 8

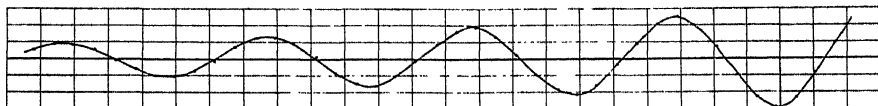


FIG. 9

15. Our tests were made with a single U-tube containing various quantities of mercury, i.e. on mercury columns of different lengths but constant diameter (0.80 cm.). In Table II the measured period T and logarithmic decrement δ , as defined in § 6, are related with the length ($2h$) of the fluid column, calculated from the weight of mercury and from the measured diameter. In every experiment the periods of successive swings were found to be practically constant, and the first four excursions, plotted logarithmi-

cally as a means to the determination of δ , gave very good straight lines; as the amplitude of excursion became smaller, δ tended to become somewhat larger.

According to the approximate theory of § 6, the period is given by

$$T = 2\pi \sqrt{\frac{h}{g'}} \times \left(1 + \frac{1}{\bar{z}\sqrt{2}}\right)$$

and the logarithmic decrement by

$$\delta = \frac{\pi}{\bar{z}\sqrt{2}} = \frac{\pi}{a} \sqrt{\frac{h\mu^2}{4g'\rho^2}}, \quad (16 \text{ bis})$$

so that for these experiments (in which $g' = g$) we have according to theory

$$T = 2\pi \sqrt{\frac{h}{g}} = T_1 \text{ (say)} \quad \left. \vphantom{\begin{matrix} T \\ T \end{matrix}} \right\} \quad (27)$$

and

$$T = T_1 \left(1 + \frac{\delta}{\pi}\right) = T_2 \text{ (say)} \quad \left. \vphantom{\begin{matrix} T \\ T \end{matrix}} \right\}$$

to a first and second approximation, while for different column lengths of the same fluid tested in a given tube, equation (16) predicts that

$$\delta^2 = \frac{\pi\mu T_1}{4\rho a^2}. \quad (28)$$

In the experiments surface tension must have operated in some manner as an additional control, and on this account we should expect the measured period to be shorter than the theoretical (since in effect g is increased). Table II shows that T is always less than T_2 , though greater than T_1 as regards the larger values of $2h$.

TABLE II

(1)	(2)	(3)	(4)	(5)	(6)	(7)
$2h$ as measured cm.	T sec.	δ (from experiments)	T_1 sec.	T_2 sec.	T/T_2	$\delta^2/T_1 \times 10^3$
22.1 ₂	0.64 ₅	0.138	0.667	0.696	0.927	2.855
44.2	0.93 ₅	0.1021 ₅	0.943	0.974	0.960	1.106
66.4	1.16	0.1124 ₅	1.156	1.197	0.969	1.094
88.5	1.36	0.1161	1.334	1.383	0.983	1.010
110.6	1.53	0.1211 ₅	1.492	1.550	0.987	0.984
132.7	1.69 ₅	0.1290	1.634	1.701	0.996	1.018

Values of $100\delta^2/T_1$ are given in the last column of Table II. They may be compared with the theoretical value according to (28), which is about

0.0058 when $\mu = 0.016$, $\rho = 13.55$. The actual decrement is much larger than our approximate theory would predict, but except for the shortest column length δ^2/T_1 is reasonably constant.

Experiments using forced oscillations

16. Our first experiments of this kind, in which glass U-tubes were employed, do not call for detailed description. They indicated the feasibility of a "forced oscillation method" of viscometry (§ 13), but they also indicated the practical desirability of the straight-tube arrangement shown in fig. 1—mainly by revealing an appreciable meniscus effect. In a bent U-tube it is difficult to estimate the effective length of the fluid column, and the circular cross-section may be altered in the bending process by an amount very difficult to measure: when a straight tube is employed, discharging into containers of greater diameter, the meniscus effect is largely reduced, the column length remains practically constant, and a high degree of circularity can be ensured. We decided to adopt this arrangement, and started experiments to determine the equivalent length of the fluid column (having regard to the accelerations at entry and discharge) and to investigate whether limitations are imposed by turbulence (§ 11).

17. Some account should be given here of the pump used in our experiments. Since our theory presumes that the pressure varies as a simple harmonic function of the time, we took pains to satisfy this condition in our apparatus by arranging to have a piston of uniform diameter moved by an eccentric rather than by a crank and connecting rod. Equally important is the requirement that excursions shall repeat indefinitely, without any tendency towards a "shifting zero": this means that on the average (i.e. in a complete cycle) the pump must move air neither into nor out of the space above the two fluid columns. To meet these requirements the pump shown (diagrammatically) in fig. 10 was constructed by Mr S. Munday in the workshops of the Engineering Department.

An annular piston (*A*) is moved up and down in a cylinder (*B*) by an eccentric (*C*) which is driven (through suitable gearing) by a small synchronous motor. The cylinder is contained within the pump body (*D*), of which the lower part (*E*) is occupied by mercury; the two U-tubes of the experiment are connected with the closed air-space (*F*) between the cylinder and pump body. The lower end of the piston is in contact with the mercury at all times: when it moves down, the displaced mercury travels up the central hole (*G*) and (since this hole is small) its free surface in *G* is considerably higher than its surface in the annular region *F*. Thus travel

of the piston creates a pressure in F , which pressure is transmitted to the U-tubes. The piston is an easy working fit in the cylinder, but surface tension prevents the mercury from leaking through the small clearance between their surfaces, and thus the mercury forms a seal preventing air from entering or leaving the air-space F .

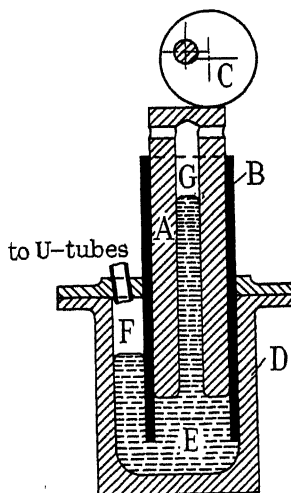


FIG. 10

18. Accurate measurement of amplitude being important, in these experiments we inserted ancillary tubes, very nearly straight and horizontal, in series with the U-tubes (fig. 7). Each contained a small quantity of coloured fluid (water or benzene), which indicated by its travel the amount of fluid displaced in the oscillation. Since the viscosity of the indicating fluid was small, very little frictional resistance was opposed; and since the tubes were only slightly inclined to the horizontal, no appreciable control was exerted by gravity.* Therefore no appreciable change of pressure occurred in the air spaces A , A , and hence no appreciable change of volume, so the indications were practically exact; and on account of their small bore the ancillary tubes magnified the excursion which had to be measured.†

19. Four steel tubes of high accuracy (hereafter termed Tubes A , B , C and D) were supplied by Messrs Accles and Pollock; their internal diameters were 0.443, 0.358, 0.153 and 0.062 in. ($a = 0.562, 0.455, 0.194, 0.078$ cm.),

* To obviate unwanted control being exerted by the inertia of the indicating column, this must be kept short.

† This device was used by one of us (A. G.) in an earlier investigation, "Viscosität und Fließfestigkeit Zaher Mineralöle" (*Berliner Akad. Berichte*, 1932).

and all four were of length 36 in. These discharged into containers having an internal diameter 1.501 (± 0.003 in.). In successive series of experiments the tube length was reduced from 36 in. to 27, 18, 12 and 6 in., and with each length different fluids were tested, having viscosities appropriate to the diameter of the tube concerned. The range of the experiments is shown in Table III. Ancillary tubes (§ 18) were used as indicators of the amplitudes of oscillation.

TABLE III

(a) denotes Castor oil (order of viscosity $\mu = 10$ poises), (b) " " "Mobiloil BB" (" " " 5 "), (c) " " "Apiezon" oil (" " " 3 "), (d) " " "Benzole" oil (" " " 2 "), (e) " " a transformer oil (" " " 0.3 "), (f) " " a fuel oil (" " " 0.1 "), (g) " " Petrol (" " " 0.01 ").					
(The numbers in the table are values of hn^2/g' , where $g' (= \pi a^2 g/A)$ is the effective value of g .)					
	Tube length (in.)				
Tube	36	27	18	12	6
<i>A</i>	10.28 (<i>a</i>), (<i>b</i>), (<i>c</i>)	7.60 (<i>a</i>), (<i>b</i>), (<i>c</i>)	5.06 ₅ (<i>a</i>), (<i>b</i>), (<i>c</i>)	3.38 (<i>a</i>), (<i>b</i>)	1.69 (<i>a</i>)
<i>B</i>	15.5 (<i>a</i>), (<i>b</i>), (<i>c</i>), (<i>d</i>), (<i>h</i>)	11.62 (<i>a</i>), (<i>b</i>), (<i>c</i>), (<i>d</i>)	7.65 (<i>a</i>), (<i>b</i>), (<i>c</i>), (<i>d</i>)	5.16 ₅ (<i>a</i>), (<i>b</i>), (<i>c</i>), (<i>d</i>)	2.38 (<i>a</i>), (<i>b</i>)
<i>C</i>	84.8 ₅ (<i>d</i>), (<i>e</i>), (<i>f</i>), (<i>h</i>)	63.6 ₅ (<i>d</i>), (<i>e</i>), (<i>f</i>)	42.4 ₅ (<i>d</i>), (<i>e</i>), (<i>f</i>)	28.3 (<i>d</i>), (<i>e</i>), (<i>f</i>)	14.15 (<i>d</i>), (<i>e</i>)
<i>D</i>	520 (<i>f</i>), (<i>g</i>)	390 (<i>f</i>), (<i>g</i>)	260 (<i>f</i>), (<i>g</i>)	173 (<i>f</i>), (<i>g</i>)	86.6 (<i>f</i>), (<i>g</i>)

* The frequency of the pressure oscillations produced by the pump, as determined from a number of observations of eccentric revolutions timed by a stop-watch, was 0.497 oscillation per second ($n = 0.98\pi = 3.079$; $n^2 = 9.480$).

In each experiment the amplitudes were observed (i) of a mercury column in a glass U-tube (the length of this column was not altered) and (ii) of the fluid in the straight tube under test. According to theory the amplitude (i) is given by (22), i.e. (hn^2/g' being constant) it is proportional to the applied pressure; the amplitude (ii) is given by (25), i.e. it is proportional to the applied pressure and inversely proportional to the effective length (h) of the tube. Therefore the ratio (response of mercury) \div (response of test fluid) should be directly proportional to h . By varying the throw of the pump eccentric (§ 17) this ratio could be determined for several amplitudes; and

by plotting mean values on a base representing the *actual* length of tube, points could be obtained which according to theory should fall on a straight line cutting this base at the point ($h = 0$). In this way the effective length was determined.

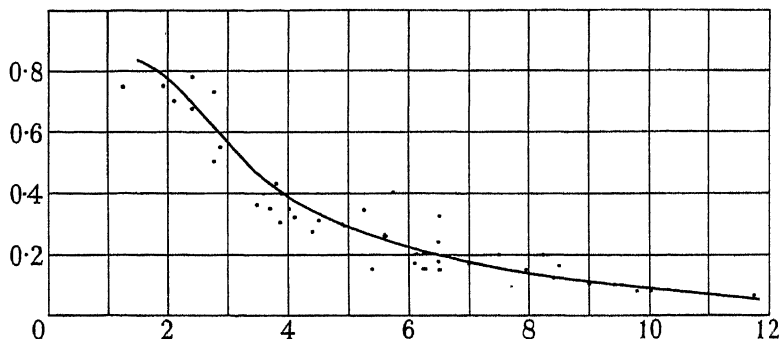


FIG. 11. Abscissae denote measured excursion of mercury indicator (in cm.). Ordinates denote addition (in cm.) to be made to measured excursion.

20. The meniscus effect in the mercury U-tube cannot be ignored when the amplitude (and therefore the restoring force of gravity) is small, but should be negligible in relation to large amplitudes. Thus it was to be expected that the ratio of the measured responses would be practically constant in experiments involving large amplitudes,* and more especially in those which employed tubes of large diameter (A and B), so that amplitudes could be measured with precision. This expectation was confirmed, and by tentatively attributing every discrepancy from a linear law to meniscus effect in the mercury it was found that a reasonably consistent correction could be deduced for that effect, in the form of an addition to the observed amplitude as given by the curve of fig. 11.

Applying the correction thus determined, we found that it considerably improved the consistency of our results: for tubes A and B the corrected ratios (Table IV) are practically independent of amplitude. From every set of experiments a limiting value, to which the ratio tends as the excursion tends to zero, can be estimated with considerable precision, and the limiting values, plotted on a base of actual length of test tube, fall very closely on straight lines (indicating a length correction independent of the actual length). The effective lengths are in all cases slightly greater (as we should expect) than the actual lengths, and the addition to be made to the actual length appears to be a constant, i.e. it is independent of that length, of the

* Meniscus effects on the fluid under test were regarded as negligible (cf. § 16).

viscosity of the fluid under test, and of the diameter. For the four sizes tested we obtained the following related values ("correction" = effective *minus* actual length):

Diameter (in.)	0.062	0.153	0.358	0.443
Correction	1.1	1.0	1.2	0.9

These figures give no indication of any systematic variation.

TABLE IV

(1) *Castor oil*

Tube A			Tube B		
Length in.	Oil amplitude cm.	Mercury Oil corrected	Length in.	Oil amplitude cm.	Mercury Oil corrected
36	2.3	1.94	36	1.0	4.4
36	3.3	1.94	36	1.5	4.3
36	4.2	1.92	36	2.6	4.35
36	5.3	1.94	36	2.8	4.50
36	6.4	1.92	36	3.4	4.32
27	3.35	1.41	27	2.1	3.4
27	6.10	1.40	27	2.75	3.4
27	6.60	1.40	27	3.00	3.45
27	8.00	1.40	27	4.55	3.40
27	11.9	1.38			
18	4.85	0.91	18	2.7	2.25
18	8.30	0.865	18	4.25	2.29
18	11.10	0.830	18	4.50	2.30
18	11.50	0.845	18	5.10	2.31
18	15.7	0.845	18	7.20	2.25
12	4.65	0.69	12	3.9	1.48
12	5.60	0.69	12	5.85	1.51
12	10.40	0.675	12	8.30	1.55
12	12.05	0.655	12	8.90	1.57
12	14.60	0.665			
6	5.5	0.355	6	4.0	0.84
6	13.3	0.347	6	8.25	0.86
6	19.0	0.343	6	11.20	0.85
6	21.6	0.340	6	12.05	0.87
			6	17.35	0.875

TABLE IV (*continued*)(2) *Mobiloil BB*

Tube A			Tube B		
Length in.	Oil amplitude cm.	Mercury	Length in.	Oil amplitude cm.	Mercury
		Oil corrected			Oil corrected
36	6.5	0.985	36	1.3	2.4
36	7.4	0.985	36	2.2	2.36
36	9.8	0.970	36	3.9	2.38
36	11.4	0.970	36	4.6	2.38
36	12.4	0.970	36	7.3	2.38
27	5.4	0.72	27	3.75	1.79
27	8.85	0.72	27	4.95	1.72
27	12.0	0.72 ₅	27	6.80	1.73
27	15.8	0.71	27	7.75	1.77
27	18.15	0.71	27	8.8	1.75
18	8.8	0.485	18	4.8	1.21
18	12.0	0.472	18	5.7	1.18
18	13.6	0.480	18	7.1	1.18
18	16.15	0.480	18	9.1	1.18
18	18.20	0.467	18	9.3	1.18
12	12.6	0.333	12	7.6	0.77 ₅
12	15.8	0.330	12	8.9	0.76 ₅
12	18.1	0.328	12	11.5	0.77 ₅
			12	12.9	0.76
			12	14.8	0.76
			12	18.2	0.77 ₅
			6	7.9	0.45 ₇
			6	9.7	0.46 ₅
			6	15.0	0.44 ₇
			6	19.0	0.45 ₃

(3) *Apiezon oil*

36	6.7	0.51 ₈	36	4.9	1.28
36	11.1	0.53 ₂	36	6.7	1.26
36	12.7	0.52 ₈	36	8.2	1.28
36	16.2	0.53 ₇	36	13.6	1.28
27	7.4	0.39 ₂	27	4.9	1.02
27	10.7	0.40 ₃	27	6.85	1.02
27	13.6	0.38 ₄	27	12.2	1.03
27	17.4	0.40 ₅	27	14.8	1.02
			27	17.1	1.02
18	10.6	0.27 ₄	18	5.45	0.65
18	11.7	0.26 ₉	18	7.45	0.64 ₅
18	16.2	0.26 ₉	18	11.4	0.65
18	18.1	0.27 ₂	18	15.8	0.66
			12	9.2	0.44
			12	17.0	0.44
			12	13.4	0.44
			12	23.2	0.44

TABLE IV (*continued*)

(4) *Benzole oil*

Tube B			Tube C		
Length in.	Oil amplitude cm.	Mercury	Length in.	Oil amplitude cm.	Mercury
		Oil corrected			Oil corrected
36	4.4	0.75	36	0.55	26.2
36	5.7	0.75	36	0.63	26.2
36	11.5	0.75	36	0.85	26
36	13.1	0.73 ₅			
36	15.4	0.74			
27	6.8	0.60 ₅	27	0.50	20.0
27	9.15	0.61	27	0.77	19.0
27	15.8	0.59 ₅	27	0.95	19.8
27	20.8	0.60			
18	13.8	0.37	18	0.76	13.2
18	14.0	0.37	18	0.96	12.6
18	15.1	0.37 ₇	18	1.41	13.2
18	20.8	0.37 ₅	18	1.24	12.7
12	12.4	0.25 ₄	12	1.42	8.7 ₅
12	15.0	0.25 ₄	12	1.65	8.4
12	16.5	0.24 ₅	12	1.80	8.6
12	21.5	0.26			
			6	1.8	5.0
			6	2.45	4.9
			6	3.3	4.9 ₅

(5) *Fuel oil*

Tube C			Tube D		
Length in.	Oil amplitude cm.	Mercury	Length in.	Oil amplitude cm.	Mercury
		Oil corrected			Oil corrected
36	4.9	0.86	36	0.34	39.5
36	7.3	0.83	36	0.43	39
36	10.5	0.85			
36	12.6	0.88			
36	20.2	0.88			
27	9.9	0.65 ₇	27	0.38	34
27	11.0	0.66 ₈	27	0.45	35
27	11.9	0.68	27	0.55	33
27	19.8	0.68			
18	10.5	0.48 ₂	18	0.44	24
18	12.7	0.48	18	0.65	24
18	15.4	0.46 ₅	18	0.78	24.8
18	19.6	0.48			
12	13.0	0.33 ₄	12	0.52	15.6
12	15.9	0.33 ₃	12	0.80	16.8
12	17.6	0.33 ₅	12	1.03	16.7
12	23.0	0.34 ₃			
			6	1.23	7.8
			6	1.58	7.7 ₅
			6	1.95	8.4 ₅
			6	2.25	8.5 ₅

TABLE IV (*continued*)

(6) <i>Transformer oil</i>			(7) <i>Petrol</i>		
Tube <i>C</i>			Tube <i>D</i>		
Length in.	Oil amplitude cm.	Mercury Oil corrected	Length in.	Oil amplitude cm.	Mercury Oil corrected
36	2.7	4.0 ₃	36	2.05	3.3 ₆
36	3.3	3.9 ₇	36	3.8	3.5 ₂
36	4.1	4.0 ₄	36	4.0	3.6 ₃
36	4.7	4.1 ₅	36	5.5	4.0
27	2.9	3.0 ₆	27	2.55	2.6 ₆
27	4.35	3.1 ₅	27	2.75	2.6 ₆
27	4.9	3.1 ₆	27	3.6	2.8
27	5.5	3.0 ₅	27	4.8	3.0
18	2.15	2.1	18	3.05	1.7 ₂
18	4.3	2.0 ₅	18	3.80	1.7 ₈
18	5.25	2.1 ₁	18	4.75	2.1 ₂
18	6.0	2.0 ₉	18	5.25	2.0 ₈
12	4.2 ₅	1.3 ₄	12	3.8	1.5 ₄
12	6.8	1.3 ₈	12	4.5	1.3 ₈
12	9.3	1.46	12	6.5	1.70
12	10.5	1.4 ₇	12	8.5	1.9 ₆
			6	3.2	0.66
			6	5.2	0.71
			6	7.8	0.76
			6	10.4	0.96
			6	12.7	1.16
			6	13.3	1.33

21. Adopting (provisionally) a length correction of 1 in. for every tube, we analysed the figures of Table IV in accordance with the formulae (22) and (25) of Part I, and in this way obtained the following estimates of viscosity (in poises):

TABLE V

Oil	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Castor oil	9.18	9.08	—	—
Mobiloil BB	4.71	4.74	—	—
Apiezon oil	2.50	2.59	—	—
Benzole oil	—	1.48	1.71	—
Transformer oil	—	—	0.275	—
Fuel oil	—	—	0.061	0.071
Petrol	—	—	—	0.0062

Exact figures not being available for comparison, the significant feature of Table V is the consistency of results obtained for the same fluid in

different tubes. This is very close as regards comparisons involving tubes *A* and *B*, less close in the comparison of *B* with *C*, and less close again in the comparison of *C* with *D*. It was to be expected that tubes of small diameter would give less consistent results, because (i) less accuracy of bore is attainable and (ii) measurement of amplitude is less easy (since the indicators have less magnifying effect). But even so Table V suggests that the theoretical formulae are reasonably correct, and its figures may be accepted as fair approximations. They are needed in examining the question whether the length correction of § 20 can be explained.

22. In this connexion two distinct effects must be recognized. First, resistance will be opposed on account of losses at entry and exit: there is a sudden change of section at each end of the tube, and each will entail some loss of "head". At velocities and for dimensions which occur in practical hydraulics, such losses vary with the squares of the relative velocities; at very low speeds (when viscosity predominates) they will vary linearly; and at intermediate speeds they may be expected to obey some intermediate (vl/ν) law. Secondly, at some more or less critical value of Reynolds' number the flow may be expected to become turbulent throughout the whole or greater part of the tube: then the whole basis of our theoretical formulae will disappear, so that measurements can be interpreted only on an empirical basis—useless for practical viscometry.

Suppose that the motion of the fluid in the reservoirs is approximately "potential flow" except at points very close to the ends of the tube. Then the loss of head due to viscosity will be inappreciable except at such points, and the reservoirs may be regarded as effectively of infinite volume. Some part of the pressure difference will be expended in accelerating the fluid from rest in the reservoirs to its full velocity in the tube, and this part will not be recovered at exit (where the fluid is decelerated) because, on account of viscosity operating on intense velocity-gradients at the ends, the motion is not reversible. There will be a "loss of head due to sudden change of section", which as stated above is known to be closely proportional to v^2 : so the pressure difference thus demanded will be proportional to ρv^2 , i.e. to

$$\rho n^2 \left(\frac{g}{g'} \right)^2 |X|^2.$$

Now the pressure difference required to maintain a response $|X|$ against the friction loss in a tube of length $2h$ and radius a is given, for low velocities, by

$$P = 16\mu h n \frac{g}{g'} |X|/a^2$$

according to (25). Therefore an "entry and exit" loss of the type considered here may be expected to involve a fractional error in the measured ratio of the responses of the mercury and test fluid, compared with its limiting value as the amplitude $\rightarrow 0$, which will be proportional to

$$\rho n^2 \left(\frac{g}{g'} \right)^2 |X|^2 / 16 \frac{\mu h}{a^2} n \frac{g}{g'} |X| = \frac{\rho n g}{16 g'} |X| \frac{a^2}{\mu h} = \frac{1}{16} R \frac{a}{h} \quad (29)$$

(R being Reynolds' number defined as in § 11) for values of R which on the one hand are not so low that the loss of head in the reservoirs is a linear function of the velocity, and on the other hand are not so high that turbulence supervenes.

As remarked in § 20, corrected for meniscus effect in the mercury the ratio of responses tends to a definite limit for every set of experiments as the amplitude of oscillation is reduced. By expressing the ratio in each experiment as a percentage of the appropriate limiting value, all of the tests can be compared on a common basis; and by fixing attention on the percentage error, the effects in question are magnified. Systematic tendency to turbulence will be revealed by departure of the results from linearity as predicted by (29).

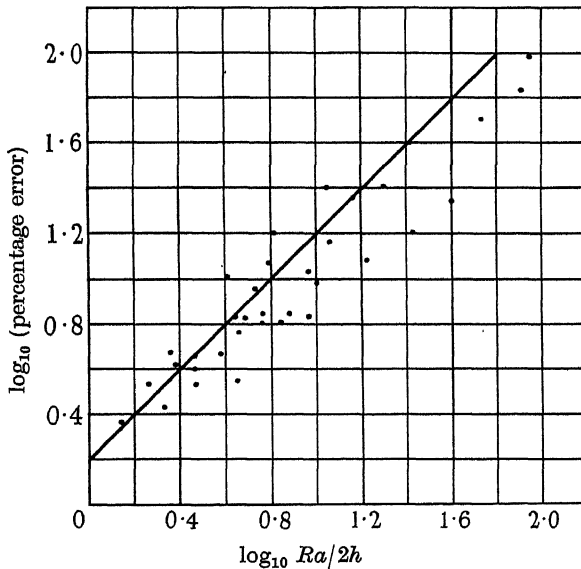


FIG. 12

23. Fig. 12 shows the experimental results plotted in this manner to logarithmic scales. According to (29) they should fall on a straight line

equally inclined to both axes, up to the point at which turbulence introduces additional losses in the tube itself. The scatter of the points is considerable, especially at the higher values of $Ra/2h$, where they relate to fluids of very low viscosity (petrol and fuel oil) tested in tubes of small diameter (difficult to measure with accuracy) and short, so that end effects have increased importance. They would have seemed more consistent if the response had been plotted as ordinate instead of the percentage error. But in a general way they confirm the prediction of (29), and suggest that a method using small excursions and fairly long tubes should give results to which, with very fair accuracy, theory can be applied.

No clear indication is afforded regarding turbulence. This is not surprising in the light of the somewhat scanty evidence which is available regarding the criterion of *steady* flow in a pipe, from which it appears that the critical value of R may be 70,000, or even higher, in the absence of imposed disturbances, and is not less than 1000* even when disturbances are present;† for in our experiments the highest value of R was about 2400, and the relevant points (in the upper range of $Ra/2h$ in fig. 12) relate as stated above to short tubes, in which end effects must be expected to mask the effects of turbulence in the tube itself. It seems safe to conclude from our results that no trouble need be expected on account of turbulence provided that $R \gg 1500$.

SUMMARY

This paper is concerned with oscillatory motion (either free or forced) of a viscous fluid along a circular tube: the restoring force comes in both instances from the "gravity head" which results from the passage of fluid. Part I is a theoretical discussion of the problem. Part II deals from a practical standpoint with complicating factors (e.g. turbulence, or meniscus and other end effects) which the theory does not take into account, and briefly discusses a possible application to viscometry. Part III gives the results of experiments made to test the predictions of Parts I and II.

REFERENCES

- Barr, G. 1931 "A Monograph of Viscometry". Oxford Univ. Press.
 Fage, A. 1938 *Proc. Roy. Soc. A*, 165, 501-29.

* In this paper we have taken a as the representative linear dimension in defining R . In most published work the diameter is taken instead, so that the numerical value of R is doubled.

† (Added 13 August 1938.) The recent experiments of Fage (1938) all related to *annular* pipes.

Jahnke and Emde 1909 "Funktionentafeln", pp. 144-7.

Kelvin, Lord 1890 "Mathematical and Physical Papers", 3, 492-3. Camb. Univ. Press.

Menneret 1911 *J. Phys. théor. appl.* 1, 753.

Poiseuille 1840-1 *C.R. Acad. Sci., Paris*, 11, 961, 1041 and 12, 112.

Russell, A. 1914 "A Treatise on the Theory of Alternating Currents", 2nd ed. 1, 210. Camb. Univ. Press.

Watson, G. N. 1923 "Theory of Bessel Functions", p. 81. Camb. Univ. Press.

The low-temperature properties of gaseous helium

BY H. S. W. MASSEY, PH.D., AND R. A. BUCKINGHAM, PH.D.

Queen's University, Belfast

(Communicated by J. E. Lennard-Jones, F.R.S.—Received 27 June 1938)

Helium is unique in that it persists as a monatomic gas down to temperatures so low that quantum effects become important. These effects arise in two more or less distinct ways. Quantum mechanics must be used in dealing with the relative motion of the gas atoms, leading to modification of the classical formulae even if the Boltzmann statistics are employed in describing the behaviour of the atomic assembly. It is further necessary, however, to allow for the symmetry properties of the atoms by making use of the Bose-Einstein instead of the classical statistics. Usually the quantization of the relative motion introduces the more important deviations from classical theory, but it is of interest to examine also whether it is possible to show, by comparison with experiment, that the Bose-Einstein statistics are the correct ones to use in dealing with helium atoms. In this paper we take account of both these quantum effects in determining theoretical values for the viscosity and second virial coefficients of helium, in the temperature range 0-25° K. Experimental values of these coefficients have been obtained at a number of temperatures in this range; comparison of observed and calculated values then leads to interesting conclusions concerning the interaction of helium atoms, the validity of the quantum theory at these low temperatures and the correct statistics to apply. At the same time results

are obtained for the total collision area of helium atoms capable of experimental tests by molecular ray methods.

The classical limit of the viscosity formula has been discussed by Massey and Mohr (1933) and by Uhlenbeck and Uehling (1933), and of that for the second virial coefficient by Uhlenbeck and Beth (1936). No classical analogue of the total cross-section exists (Massey and Mohr 1933). In an accompanying paper by one of us (R. A. B.) the classical theory of the equation of state of helium and also of neon and argon is discussed and the conclusions drawn therefrom used to supplement those of this paper in providing information as to the form of the interaction of helium atoms. A complete knowledge of this interaction may be important for a discussion of the remarkable properties of liquid helium.

1. THE QUANTAL FORMULAE

Let $V(r)$ be the energy of interaction between two helium atoms at distance r apart. The wave equation for the relative motion of the two atoms, each of mass M , with energy $E = k^2\hbar^2/4\pi^2M$ and angular momentum $\sqrt{l(l+1)} \frac{\hbar}{2\pi}$ is

$$\frac{d^2}{dr^2}(r\psi_l) + \left(k^2 - \frac{4\pi^2M}{\hbar^2}V - \frac{l(l+1)}{r^2}\right)(r\psi_l) = 0. \quad (1)$$

The asymptotic form of the solution of this equation which is a proper function at the origin is

$$r^{-1} \sin(kr - \frac{1}{2}l\pi + \delta_l).$$

It is possible to express the formulae for the total elastic cross-section, viscosity coefficient and second virial coefficient in terms of the phases δ_l as follows.

The total elastic cross-section for collision between two gas atoms with relative energy E is (Mott and Massey 1933)

$$Q_0(E) = \frac{4\pi}{k^2} \sum (2l+1) \sin^2 \delta_l. \quad (2)$$

If no symmetry properties need be taken into account the summation is taken over all integral values of l from 0 to ∞ , but if the Bose-Einstein statistics are valid the sum is twice that over even values of l only, while in the Fermi-Dirac case it is twice that over odd values only.

The coefficient of viscosity $\eta(T)$ at temperature T is given in terms of the viscosity cross-section $Q_\eta(E)$ by Chapman's formula (1917):

$$\eta(T) = \frac{10\kappa^3 T^3}{M^2} \left(\frac{4\pi\kappa T}{M} \right)^{\frac{1}{2}} \frac{1}{\pi R}, \quad (3)$$

where

$$R = \frac{1}{2} \left(\frac{h}{M\pi} \right)^8 \int_0^\infty Q_\eta k^7 e^{-k^2/k_0^2} dk,$$

$$k_0 = \frac{2\pi}{h} (M\kappa T)^{\frac{1}{2}}.$$

$Q_\eta(E)$ is then given by (Massey and Mohr 1933)

$$Q_\eta(E) = \frac{4\pi}{k^2} \sum \left\{ \frac{4l^3 + 6l^2 - 2l - 2}{(2l-2)(2l+3)} \sin^2 \delta_l - \frac{2(l+2)(l+1)}{2l+3} \right. \\ \left. \times \cos(\delta_l - \delta_{l+2}) \sin \delta_l \sin \delta_{l+2} \right\}, \quad (4)$$

the limits for the summation being the same as for $Q_0(E)$.

For the second virial coefficient $B(T)$ we have the formula

$$B(T) = -\frac{Nh^3}{16(\pi M\kappa T)^{\frac{3}{2}}} \left\{ \epsilon + \frac{8}{\pi} \sum_l (2l+1) \int_0^\infty \frac{d\delta_l}{dk} e^{-k^2/k_0^2} dk \right\} + B_{\text{disc}}, \quad (5)$$

given independently by Gropper (1937) and Beth and Uhlenbeck (1937). With Boltzmann statistics $\epsilon = 0$, and the summation is over all integral values of l from 0 to ∞ . For the Bose-Einstein case $\epsilon = 1$, and the summation is twice that over even values of l , while in the Fermi-Dirac case $\epsilon = -1$ and twice the sum over odd values of l is to be taken.

B_{disc} represents the contribution from discrete quantum states for the relative motion, and takes the form

$$B_{\text{disc}} = -\frac{Nh^3}{2(\pi M\kappa T)^{\frac{3}{2}}} \sum_n g_n e^{E_n/\kappa T}, \quad (6)$$

where g_n is the statistical weight of the level of energy E_n . Of course only even or odd levels will contribute if Bose-Einstein or Fermi-Dirac statistics, respectively, are valid. With the interaction energies typical of gas atoms only a finite number of discrete quantum states occur, so there is no difficulty about convergence in calculating B_{disc} . The convergence of the series involving the phases depends on the fact that $\delta_l \rightarrow 0$ when $kr \sim l$ for such values of r that $\frac{4\pi^2 M}{h^2} |V(r)| \ll \frac{l(l+1)}{r^2}$ (Mott and Massey 1933). This condition

shows that, except at very low temperatures, we must expect to take into account a considerable number of terms.

2. THE ATOMIC INTERACTION AND THE CALCULATION OF PHASES AND ENERGY LEVELS

To carry out detailed calculations it is necessary to assume some form for the interaction energy $V(r)$. For this function we employed the form

$$V(r) = 7.7 \times 10^{-10} e^{-r/0.217} - 1.47 \times 10^{-12} r^{-6} \text{ ergs,} \quad (7)$$

(r in \AA)

given by Slater (1928) as likely to represent the best initial approximation.

To solve the wave equation (1) with this function one must have recourse to mechanical or numerical methods of solution, if approximate methods are not to be employed. We therefore solved the equations of form (1) which arose by means of a model differential analyser a description of which has been published elsewhere (Massey, Wylie, Buckingham and Sullivan 1938). Checks of the performance of the machine were made by numerical integration at suitably chosen values of the constants.

As the form (7) does not represent the interaction even approximately for small distances r the integration could not be started from the origin. Instead the following procedure was adopted. The equation (1) was integrated inwards from the zero r_0 of the function

$$k^2 - \frac{l(l+1)}{r^2} + \frac{4\pi^2 M}{h^2} V(r),$$

assuming different initial values of $d\psi/dr$, until a value was found which gave a solution tending asymptotically to zero at the origin. With this slope the equation was then integrated outwards from r_0 and the asymptotic form found. The sensitivity of the determination of the "proper" slope at r_0 is so high that the inner integration need not be carried so far that (7) becomes invalid.

In Table I we give the values of the phases for a number of values of k and l in the range of importance.

The calculation of the discrete energy values presented no difficulty. Only one level exists, for $l = 0$ and for it

$$E_0 = 0.0083 \times 10^{-15} \text{ erg.}$$

3. CALCULATION OF THE COEFFICIENT OF VISCOSITY

Using the formula (4) and the values of δ_i given in Table I, we obtain the viscosity cross-section Q_η as a function of E illustrated in fig. 1 on a

logarithmic scale. Curves I, II, III are obtained using Bose-Einstein, Fermi-Dirac and classical statistics respectively. A second horizontal scale is attached indicating the temperature at which the corresponding value of E gives the maximum contribution to the viscosity coefficient as represented by the formula (3). It is then clear that we can only expect appreciable differences between the results obtained with different statistics at very low temperatures indeed.

TABLE I. VALUES OF PHASES δ_l

$l \backslash ka_0$	0	1	2	4	6	8	10
0.01	2.79₇	—	—	—	—	—	—
0.05	(1.91)	0.04 ₃	—	—	—	—	—
0.10	1.23₀	0.30 ₅	0.00 ₇	—	—	—	—
0.15	(0.75)	0.75	(0.06)	—	—	—	—
0.2	0.40	0.85	0.16₄	—	—	—	—
0.3	— 0.26₂	0.60 ₅	0.43	0.01	—	—	—
0.4	(—0.84)	0.27	0.65₅	(0.08)	—	—	—
0.5	—1.37	—0.19 ₅	0.43 ₅	0.19₉	—	—	—
0.6	(—1.86)	—	(0.16)	0.33₅	—	—	—
0.667	—2.22	—	—0.06	0.41	0.10 ₃	—	—
0.889	—3.27	—	—0.92	0.26	0.20	0.06	—
1.1	—4.36	—	(—1.75)	(—0.20)	0.31	0.16 ₇	0.03 ₂
1.333	—5.32	—	—2.67	—0.80 ₅	0.18 ₇	0.26	0.09 ₃
2.0	—7.83	—	—5.07	—2.84	—1.23	—0.17	0.21

Phases in heavy type have been calculated by accurate numerical integration, and the remainder (except those in brackets, which have been interpolated) by mechanical integration of the wave equation.

The viscosity coefficient may be obtained as a function of temperature from (3) by numerical integration. Values obtained in this way are given in Table II and compared with the recent experimental values of van Itterbeek and Keesom (1938). The agreement at temperatures above 15° K is quite good, but the calculated values are slightly too large, that is to say the theoretical viscosity cross-section is too small. Below 15° K the agreement becomes less satisfactory and the discrepancy is in the opposite sense. This will be discussed further in § 6 along with the results derived from calculation of the second virial coefficient. At the lowest temperatures the values obtained neglecting symmetry properties are also given (column II), and it is again clear that no conclusions can be drawn regarding the correct statistics to employ, though the calculated values of column I have been derived on the reasonable assumption of the validity of the Bose-Einstein statistics.

TABLE II. VISCOSITY OF HELIUM (IN 10^{-6} C.G.S. UNITS)

T ($^{\circ}$ K)	(Exp.)	(Calc.)		
		I	II	III
1.64	5.47	(3.4)	—	—
2.0	—	4.35	5.05	—
3.0	—	7.10	7.75	—
4.23	12.67	(10.7)	—	—
5	—	12.75	—	—
10	—	22.3	—	—
14.27	27.87	(28.4)	—	—
15.0	29.46*	29.4	—	36
16.55	30.77	(31.4)	—	—
20.2	35.03*	36.1	—	43
20.38	35.19	36.3	—	—
25	—	41.2	—	—
30	—	46.3	—	—

Experimental results of van Itterbeek and Keesom (1938) except those marked *, for which see *International Critical Tables*.

Calculated values:

- I. Using Bose-Einstein statistics. () interpolated.
- II. Without allowing for symmetry properties.
- III. Values given by Massey and Mohr (1934).

It is of interest to compare the results of the present calculation with those obtained by Massey and Mohr (1934) using the same interaction energy (7) but calculating the phases by a combination of Jeffreys' and Born's methods. They found the values given in column III of Table II which are too large, i.e. indicating that the calculated viscosity cross-sections were too small. The reason for this is that the perturbation method used to obtain the smaller (higher order) phases is not capable of giving with any accuracy the positive phase shifts due to the attractive part of the field. The contribution from the repulsive field is exaggerated in such a method, so the positive phases obtained were all too small. The same objection applies to the values obtained by Massey and Mohr by a second perturbation method in which the actual interaction function was approximated by an analytical expression for which the phases could be exactly calculated and the effect of the surplus field allowed for by a perturbation calculation. It is generally true that perturbation methods are only capable of giving accurate results, under very stringent conditions, when applied to problems involving strong repulsive fields even when the perturbation is quite small. As the higher order phases become relatively less important as the number of terms involved in the series increases, the values obtained by Massey and Mohr for higher temperatures are more accurate.

4. TOTAL ELASTIC CROSS-SECTION

It is now possible to investigate, by molecular ray methods, the free paths of gas atoms and hence the total collision cross-sections. Experiments of this type in helium would be very useful in providing additional information as to the form and magnitude of the interatomic force. The calculated form

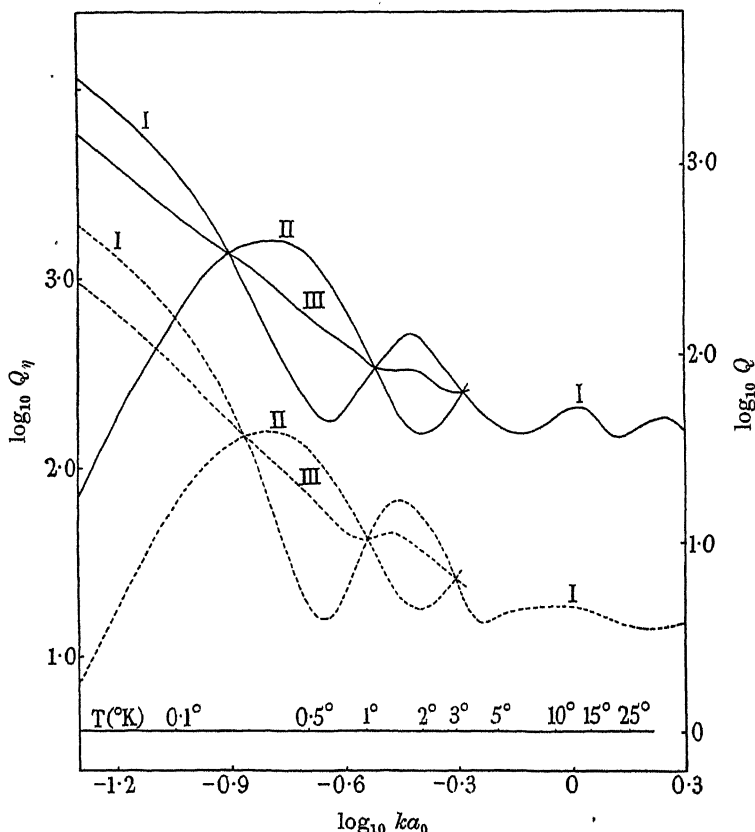


FIG. 1. Calculated cross-section for helium atoms. — Total cross-section Q . ---- Viscosity cross-section Q_η . I, Bose-Einstein statistics. II, Fermi-Dirac statistics. III, Without allowance for symmetry properties.

for the total cross-section as a function of the relative energy of the colliding atoms is given in fig. 1 on a logarithmic scale. The oscillatory behaviour of this function results mainly from the assumption of the Bose-Einstein statistics and not from a Ramsauer effect due to the attractive field. This may be seen by comparison with the curve calculated without taking account of symmetry properties.

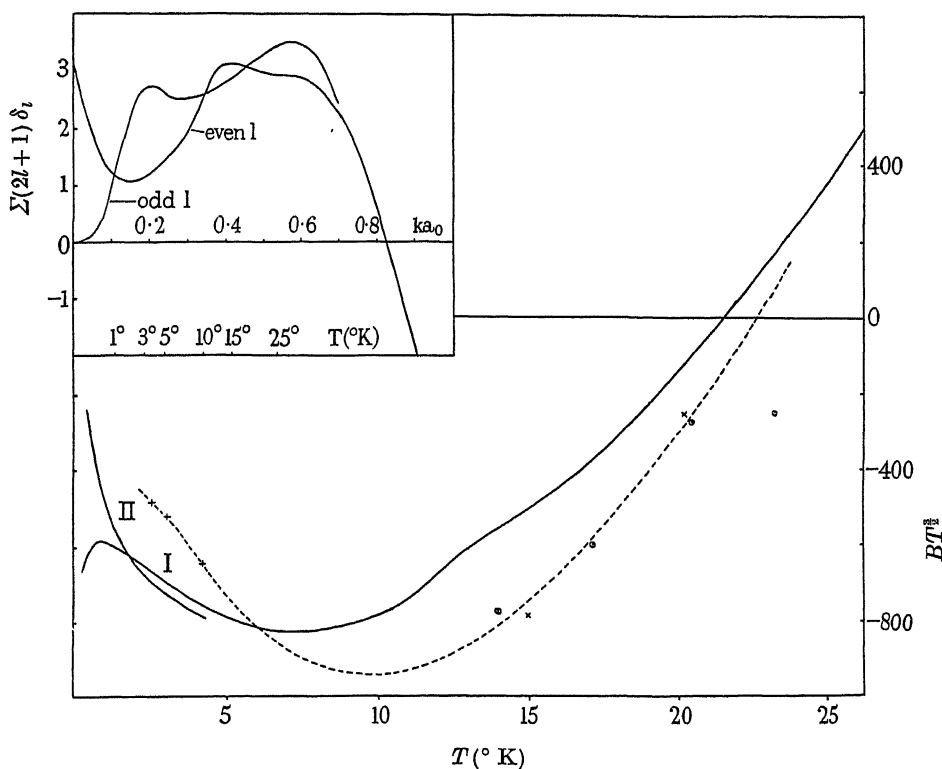


FIG. 2. Second virial coefficient of helium. I, Calculated assuming Bose-Einstein statistics. II, calculated assuming Fermi-Dirac statistics. - - - probably correct form of curve. + Keesom and Kraak. x Holborn and Otto. o Nijhoff, Iliin and Keesom.

5. THE SECOND VIRIAL COEFFICIENT

In fig. 2, curve I, illustrates the form of the function $BT^{3/2}$ between $T = 0^\circ$ and 26°K , calculated from (5), assuming Bose-Einstein statistics and the phases given in Table I. The dotted curve II represents that given by the Fermi-Dirac statistics. This curve has not been obtained with such accuracy as the first, for, of the odd order phases, δ_1 alone was calculated directly, the remainder being determined by interpolation. The third curve represents what is probably the true curve as indicated by the experimental points of Keesom and Kraak (1935), Nijhoff, Iliin and Keesom (1928) and Holborn and Otto (1924-6).

Features of the experimental results reproduced by the theory are the maximum in the neighbourhood of 10°K and the zero (Boyle point) at about 21°K , while at the same time the order of magnitude is given quite well.

The chief discrepancy is that the calculated features are displaced to too low temperatures in each case. This is clearly to be attributed to defects in the interatomic interaction assumed which were found in § 3 to lead also to incorrect values of the viscosity coefficients at low temperatures and will be discussed further in § 6. Even allowing for this inexactness of the interaction, it is abundantly clear that there is little possibility of distinguishing the correct form of statistics by means of the virial coefficient except by measurement at very low temperatures indeed ($< 1^\circ \text{K}$).

It is of some interest to trace the form of the calculated curves to the behaviour of the phases (see Table I). At very low energies of relative motion all phases but δ_0 tend to $+0$ with increasing rapidity as the order increases. δ_0 tends to a value π owing to the existence of a discrete energy level. As the energy increases the phases rise successively to positive maxima after which they decline to zero and increase rapidly in the negative direction. Consider first the behaviour of $B_B T^{\frac{3}{2}}$, where B_B refers to the value calculated on the assumption of Bose-Einstein statistics. Integrating by parts and remembering that δ_0 tends to π we have

$$B_B T^{\frac{3}{2}} = \frac{\hbar^3 N}{16(\pi M \kappa)^{\frac{3}{2}}} \left[-1 - \frac{32}{\pi k_0^2} \int_0^\infty e^{-k^2/k_0^2} \sum_l (4l+1) \delta_{2l} k dk - 16(e^{E_0/\kappa T} - 1) \right]. \quad (8)$$

At very low temperatures ($< 1^\circ \text{K}$) the effect of the discrete level predominates giving a term falling off rapidly from negative infinity with increasing temperature. Owing to the smallness of E_0 this term soon becomes small and the dominating term is the second. The sum over the phases is illustrated as a function of k in the inset fig. 2 with an appropriate temperature scale attached. The minimum at $T = 2^\circ \text{K}$ appears because of the decrease of δ_0 combined with increase of δ_2 , and this gives rise to the maximum in $B_B T^{\frac{3}{2}}$ at 1°K . As T increases, first one and then a second broad maximum occurs due to δ_2 and δ_4 . After this the sum becomes rapidly negative due to the rapid attainment of large negative values by δ_0 , δ_2 and δ_4 . The result is then to produce first the minimum in $B_B T^{\frac{3}{2}}$ at 7°K and then the rapid change of sign involving the zero at 21°K .

The behaviour of $B_F T^{\frac{3}{2}}$ for the Fermi-Dirac case is very similar. We have

$$B_F T^{\frac{3}{2}} = \frac{N \hbar^3}{16(\pi M \kappa)^{\frac{3}{2}}} \left\{ 1 - \frac{32}{\pi k_0^2} \int_0^\infty e^{-k^2/k_0^2} \sum_l (4l+1) \delta_{2l+1} k dk \right\},$$

where there is no influence from any discrete energy level. At very low temperatures ($< 1^\circ \text{K}$) this falls below $B_B T^{\frac{3}{2}}$ in absolute magnitude not only because of the absence of the exponential $e^{E_0/\kappa T}$ but also because δ_1

is less than δ_0 under these conditions. This is illustrated from the phase sum curve in the inset fig. 2. In all cases the second term dominates as the temperature increases. The phase sum rises to a maximum due to δ_1 almost in the same region where the corresponding sum for B_B is a minimum (2° K). Owing, however, to the fact that in $B_{B'}$, unlike B_B , the first term is of opposite sign to the second, the net result is to give nearly the same values for both $B_{B'}$ and B_B in this range (above 1.5° K). At higher temperatures still the difference between the two phase sums becomes small and the effect of the first term unimportant, so B_B and $B_{B'}$ tend rapidly to equality.

6. DISCUSSION OF THE RESULTS

It remains for us to discuss the results obtained from comparison of calculated and observed values for the viscosity and second virial coefficients in so far as they throw light on the interatomic interaction. We assume that the Bose-Einstein statistics are valid although, as we have seen, there is nothing in the observed results which could distinguish the various possibilities.

From § 5 we see that, in the temperature range $2\text{--}7^\circ$ K, the calculated absolute value of the second virial coefficient is greater than the observed. This would be rectified by reducing the interaction energy at large distances for, on the one hand, this would remove the discrete energy level and on the other reduce the contribution from δ_2 in this temperature range. Since, however, δ_2 must attain about as large a maximum as before although at greater energies, the attractive field at smaller distances must be somewhat increased. Rough estimates based on the behaviour of the phases seem to indicate that the field at separations greater than $3.5\text{--}5$ Å must be reduced by between 10 and 20 %. Such a modification of the field would also tend to remove the discrepancy below 10° K in the viscosity coefficient. At these temperatures the positive phases predominate in determining Q_η ; reduction in the attractive field at large distances will reduce these phases and hence Q_η , giving a larger viscosity. Estimates show that the same order of magnitude is required for the modification of field.

It is not difficult to justify these modifications theoretically. The van der Waals' force constant for helium is quite likely to be overestimated by as much as 20 % in (7), and this would give too large a field at large distances. Now in the formula (7) dipole-quadrupole interaction terms, which give a contribution varying as r^{-8} for large r , are neglected. Approximate estimates (Buckingham 1937) indicate that at the field minimum this term contributes 20–25 % of the r^{-6} term. If such a term were included in (7)

with reduced van der Waals' force constant for the r^{-6} term, the modified attractive field would be reduced at large distances while remaining practically unaltered or even being increased at smaller. This would be of the required form. The small deviations in the viscosity coefficient above 10° could easily be removed by small changes in the magnitude of the repulsive field.

Although the quantal theory worked out in this paper does not give results in exact agreement with experiment it is clear that the discrepancies can be removed by theoretically plausible modifications of the interatomic interaction energy assumed. These modifications are not incompatible with the results found in the accompanying paper by one of us (R. A. B.) on the virial coefficients at higher temperatures.

It should not be difficult to obtain an interaction function of considerable accuracy and it is hoped to do this in a later paper. An accurate field would be very valuable for theoretical developments concerning liquid and solid helium.

SUMMARY

The viscosity and second virial coefficients of helium gas are calculated in the temperature range 0 – 26°K using exact quantum formulation and the interatomic interaction given by Slater. It is found that there is very good agreement with the observed viscosity down to 15°K , but below this temperature the calculated values are too small. The correct form for the variation of the virial coefficient with temperature is obtained, but the calculated behaviour is in error in predicting too low temperatures for the position of the maxima, etc. (The calculated Boyle point is 21° compared with the observed 23° .)

It is shown that the discrepancies may be greatly reduced in both cases by small modification of the Slater field involving reduction of the van der Waals' force at large distances and inclusion of dipole-quadrupole interaction. No difference between the theoretical results obtained by assuming different statistics is sufficiently definite to be distinguished from small errors in the assumed interaction, in the temperature range from 1°K upward.

REFERENCES

- Beth, E. and Uhlenbeck, G. E. 1937 *Physica*, **4**, 915.
Buckingham, R. A. 1937 *Proc. Roy. Soc. A*, **161**, 117.
Chapman, S. 1917 *Philos. Trans. A*, **217**, 115.
Gropper, L. 1937 *Phys. Rev.* **51**, 1108.

- Holborn, L. and Otto, J. 1924-6 *Z. Phys.* **23**, 77; **30**, 320; **33**, 1; and **38**, 359.
van Itterbeek, A. and Keesom, W. H. 1938 *Physica*, **5**, 257.
Keesom, W. H. and Kraak, H. H. 1935 *Physica*, **2**, 37.
Massey, H. S. W. and Mohr, C. B. O. 1933 *Proc. Roy. Soc. A*, **141**, 434.
— — 1934 *Proc. Roy. Soc. A*, **144**, 188.
Massey, H. S. W., Wylie, J., Buckingham, R. A. and Sullivan, R. 1938 *Proc. R. Irish Acad.* (In course of publication.)
Mott, N. F. and Massey, H. S. W. 1933 "Theory of Atomic Collisions", p. 24.
Nijhoff, G. P., Iliin, B. and Keesom, W. H. 1928 *Commun. Phys. Lab. Univ. Leiden*. p. 1886; *Proc. K. Akad. Wet. Amst.* p. 408.
Slater, J. C. 1928 *Phys. Rev.* **32**, 349.
Uhlenbeck, G. E. and Beth, E. 1936 *Physica*, **3**, 729.
Uhlenbeck, G. E. and Uehling, E. 1933 *Phys. Rev.* **43**, 552.
-

The electromagnetic energy of a point charge

BY M. H. L. PRYCE, *Trinity College, Cambridge*

(Communicated by P. A. M. Dirac, F.R.S.—Received 30 June 1938)

1. INTRODUCTION

There are many reasons for preferring the point model of the electron, in which the field equations of empty space hold all the way up to the centre of the electron, to the Lorentz model, in which the charge is distributed over a small sphere. The point model is not without difficulties, however, and two have attracted special attention. The first is that the field becomes infinite at the charge, so that the Lorentz equations of motion cannot be applied directly; the second is that the ordinary expression leads to an infinite electromagnetic energy in the neighbourhood of the charge. As these difficulties occur both in classical and in quantum electrodynamics it seems reasonable to look for their solution, first in the classical theory, and then try to translate it into the quantum theory. A recent paper by Dirac (1938) has satisfactorily removed the first difficulty from the classical theory. The present paper shows how the second can be removed also. The translation of these methods to quantum theory has not yet been accomplished. Some papers by Wentzel (1933, 1934) have also dealt with this subject, both from the classical and the quantum standpoint, but they do not seem to be altogether without difficulties, and the method is rather complicated to use in actual problems.

2. THE ELECTROMAGNETIC ENERGY

The ordinary expression for the electromagnetic energy in a region of space is

$$\frac{1}{8\pi} \int (\mathbf{E}^2 + \mathbf{H}^2) dV,$$

where \mathbf{E} , \mathbf{H} are the electric and magnetic vectors, and dV is the element of volume. The classical derivation of this is briefly as follows: one calculates the rate at which the electromagnetic forces, $\rho\mathbf{E} + \mathbf{j} \times \mathbf{H}$ per unit volume (ρ , \mathbf{j} being the charge and current density), do mechanical work on the charge distributions in the region. This is $c\mathbf{j}\mathbf{E}dV$, which can be transformed by means of the Maxwell equations to

$$-\frac{1}{8\pi} \frac{\partial}{\partial t} \int (\mathbf{E}^2 + \mathbf{H}^2) dV - \frac{c}{4\pi} \int [\mathbf{E} \times \mathbf{H}] \cdot d\mathbf{S}$$

the second integral being taken over the boundary of the region. Thus the rate of change of the sum of the mechanical energy and $\int (\mathbf{E}^2 + \mathbf{H}^2) dV/8\pi$ is given by the surface integral $c\int [\mathbf{E} \times \mathbf{H}] \cdot d\mathbf{S}/4\pi$; this can be interpreted as meaning the conservation of energy if $\int (\mathbf{E}^2 + \mathbf{H}^2) dV/8\pi$ is the electromagnetic energy in the region and $\int c[\mathbf{E} \times \mathbf{H}] \cdot d\mathbf{S}/4\pi$ the rate of flow of energy out of the region across the boundary. This derivation is strictly valid only if ρ is everywhere finite (but not necessarily continuous), and therefore does not apply to point charges. Furthermore it gives only the rate of change of the energy, leaving the energy itself indeterminate to the extent of an additive constant.

It is reasonable to assume that the concentration of the charge into a point will not alter the expression for the energy in a region of space, apart from an additive constant, if this region is free from charges. (A convenient formulation of this is as follows: if two electromagnetic fields differ only in a region containing no charges, the difference of the electromagnetic energies associated with them is the integral of the difference of their energy densities $U (= [\mathbf{E}^2 + \mathbf{H}^2]/8\pi)$ over that region.)

It is possible to write down a finite expression for the energy, consistent with this principle. Let it be assumed for convenience that there is only one charge in the field; all that follows can be easily extended to many charges. Then let the charge be surrounded by a small closed surface, whose diameter is smaller than any physically important dimension and will eventually be made to tend to zero, and let the integral of the energy density U be formed over the whole of space outside the surface. As the diameter tends to zero the integral tends to infinity. If, however, a quantity

can be found, depending only on the surface and the variables describing the charge, such that its sum with the integral tends to a finite limit, then this sum will satisfy the above requirement. This will be so if a vector field \mathbf{K} , depending only on the variables of the charge, can be found such that $U' = U - \text{div } \mathbf{K}$ is integrable in the neighbourhood of the charge (i.e. does not go to infinity more rapidly than r^{-2}). A possible expression for the electromagnetic energy is then

$$W_{\text{el}} = \lim \left\{ \int U dV + \int \mathbf{K} \cdot d\mathbf{S} \right\}, \quad (2.1)$$

the volume integral being taken over the region of space outside the surface, and the surface integral being taken with the outward normal. Equation 2.1 expresses the central idea of this paper. Loosely speaking, the first term is the usual infinite energy and the second term is a negatively infinite energy arising from the charge itself.

If \mathbf{K} falls off at infinity more rapidly than R^{-2} , this can also be written as an integral over the whole of space:

$$W_{\text{el}} = \int (U - \text{div } \mathbf{K}) dV = \int U' dV. \quad (2.2)$$

Although the physical significance is hidden in this formulation, it is often more convenient to work with it than with equation 2.1.

The *total* energy is of course the sum of the electromagnetic energy and some energy associated with the charge itself, which may be called the mechanical energy. The distinction between electromagnetic and mechanical energy is not sharp, for a different vector \mathbf{K} will in general give a different value of the electromagnetic energy, which must be compensated in the mechanical energy to keep the total unaltered.

For this procedure to be physically correct two conditions must be satisfied: the conservation of energy must follow from the field equations and the equations of motion for a point charge, and it must be relativistically invariant. Momentum and energy must therefore be treated on an equal footing.

3. RELATIVISTIC FORMULATION

Relativistic vector and tensor notation will be used throughout. The velocity of light will be taken as unit. Greek indices take the values 0, 1, 2, 3; Roman indices take the values 1, 2, 3. The co-ordinates of a point in space-time are written x^μ ; x^0 refers to the time co-ordinate, x^1, x^2, x^3 to the space co-ordinates. Indices are raised and lowered by means of the metric tensor

$g_{\mu\nu}$ ($g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$, $g_{\mu\nu} = 0$ if $\mu \neq \nu$). The electromagnetic field is given by the antisymmetric tensor $F_{\mu\nu}$ ($E_1 = F_{01}$, $H_1 = F_{32}$, etc.). The total energy and momentum are components of a vector p^σ , which can be separated into an electromagnetic and a mechanical part,

$$p^\sigma = p_{\text{el}}^\sigma + p_{\text{mech}}^\sigma,$$

although, as already noted, the division is not absolute.

The ordinary expression for p_{el}^σ is given by the space integral of the component $T^{0\sigma}$ of the energy tensor:

$$p_{\text{el}}^\sigma = \int T^{0\sigma} dx^1 dx^2 dx^3,$$

with

$$T^{\nu\sigma} = \frac{1}{4\pi} \{ F^{\nu\alpha} F_\alpha^\sigma + \frac{1}{4} F^{\alpha\beta} F_{\alpha\beta} g^{\nu\sigma} \}. \quad (3.1)$$

($T^{00} = U$). $T^{\nu\sigma}$ is symmetric and satisfies the conservation law

$$\frac{\partial}{\partial x^\nu} T^{\nu\sigma} = 0. \quad (3.2)$$

In the relativistic theory, the components of \mathbf{K} , together with those of the similar (three-dimensional) vectors associated with the components of momentum, can be seen to form part of a tensor of third rank $K^{\mu\nu\sigma}$, antisymmetric in μ and ν ,

$$K^{\mu\nu\sigma} = -K^{\nu\mu\sigma}. \quad (3.3)$$

Those of \mathbf{K} are in fact K^{a00} ($a = 1, 2, 3$), and those of the vector associated with the component p_{el}^σ are $K^{a0\sigma}$. The (three-dimensional) divergence of these vectors must differ from $T^{0\sigma}$ by an integrable amount. Because of the antisymmetry, the three-dimensional divergence is the same as $\partial K^{\mu 0\sigma} / \partial x^\mu$. Thus $T'^{\nu\sigma}$, defined analogously to U' by

$$T'^{\nu\sigma} = T^{\nu\sigma} - \frac{\partial}{\partial x^\mu} K^{\mu\nu\sigma}, \quad (3.4)$$

must be integrable. $T'^{\nu\sigma}$ also satisfies a conservation law

$$\frac{\partial}{\partial x^\nu} T'^{\nu\sigma} = 0 \quad (3.5)$$

at all points not on the world-line of the charge. This follows from 3.2 and the antisymmetry of $K^{\mu\nu\sigma}$.

At first sight it would not seem necessary to require that $T'^{\nu\sigma}$ should be symmetric; but if these methods are also applied to discuss angular momentum, it will be convenient to have it so. It will appear later that $K^{\mu\nu\sigma}$ can

be so chosen that $\partial K^{\mu\nu\sigma}/\partial x^\mu$ is symmetric, although $K^{\mu\nu\sigma}$ itself is not symmetric in ν and σ . It is therefore natural to add this restriction.

The relativistic extension of equation 2.2 is

$$p_{\text{el}}^\sigma = \int T'^{0\sigma} dx^1 dx^2 dx^3 \quad (3.6)$$

This is a more convenient form to take than the extension of 2.1, for it can be discussed with less mathematical apparatus. It is of course strictly valid only if $K^{\mu\nu\sigma}$ falls off more rapidly than R^{-2} at infinity, but since it is only its behaviour near the charge that is important and at large distances $K^{\mu\nu\sigma}$ may be made to behave as we please, this is unimportant.

Let $p_{\text{el}}^\sigma(1)$ and $p_{\text{el}}^\sigma(2)$ denote the values at two times t_1 and t_2 . These are given by integrals over two three-dimensional sections of space-time, $x^0 = t_1$ and $x^0 = t_2$. Let the world-line of the charge be surrounded by a tube of small diameter (a section $x^0 = \text{const.}$ will cut the tube in a closed surface surrounding the charge); the diameter of the tube will ultimately tend to zero. In the region of space-time between the sections $x^0 = t_1$ and $x^0 = t_2$, and outside the tube, 3.5 holds exactly. From this it follows, by the theorem of Gauss and Green, that the integral of the "normal component" of $T'^{\nu\sigma}$ over the boundary of the region is zero; i.e.

$$\int_1 T'^{\nu\sigma} dS_\nu - \int_2 T'^{\nu\sigma} dS_\nu + \int_{\text{tube}} T'^{\nu\sigma} dS_\nu = 0, \quad (3.7)$$

where \int_1, \int_2 are taken over the part of the sections $x^0 = t_1$ and $x^0 = t_2$ outside the tube, and dS_ν is the directed "surface element"; in the first integral it has been taken in the inward direction to correspond with the positive direction of time (it is assumed that $t_2 > t_1$), and away from the world-line on the tube. As the diameter of the tube tends to zero, the first two integrals tend to $p_{\text{el}}^\sigma(1)$, $p_{\text{el}}^\sigma(2)$ respectively. The integral over the tube therefore tends to a definite limit, irrespective of the shape of the tube. This limit must be expressible as a line integral along the world-line. If τ denotes the proper time along the world-line, measured from some arbitrary zero, it may be written

$$\lim \int_{\text{tube}} T'^{\nu\sigma} dS_\nu = - \int_1^2 R^\sigma d\tau. \quad (3.8)$$

Equation 3.7 therefore gives

$$p_{\text{el}}^\sigma(2) - p_{\text{el}}^\sigma(1) = - \int_1^2 R^\sigma d\tau. \quad (3.9)$$

Energy and momentum will be conserved if p_{mech}^σ satisfies the equation of motion

$$\dot{p}_{\text{mech}}^\sigma = R^\sigma. \quad (3.10)$$

(A dot means differentiation with respect to the proper time.)

If a different $K^{\mu\nu\sigma}$ is taken, differing from the first by $\Delta K^{\mu\nu\sigma}$ ($\partial \Delta K^{\mu\nu\sigma} / \partial x^\mu$ being integrable), the new p_{el}^σ will differ from the old by an amount

$$\Delta p_{\text{el}}^\sigma = \lim \int \Delta K^{a0\sigma} d\sigma_a,$$

where $d\sigma_a$ stand for the components of the surface element of the surface round the charge, in the three-dimensional space $x^0 = \text{const}$. This limit exists, for $\partial \Delta K^{a0\sigma} / \partial x^a$ is integrable. Equation 3.9 shows that R^σ is changed by an amount

$$\Delta R^\sigma = -\Delta \dot{p}_{\text{el}}^\sigma.$$

The change in p_{mech}^σ compensates the change in p_{el}^σ :

$$\Delta p_{\text{mech}}^\sigma = -\Delta p_{\text{el}}^\sigma.$$

The equation of motion now reads

$$\dot{p}_{\text{mech}}^\sigma + \Delta \dot{p}_{\text{mech}}^\sigma = R^\sigma + \Delta R^\sigma,$$

which is of course identical with 3.10.

Equation 3.7 still holds if the sections $x^0 = t_1$, $x^0 = t_2$ are replaced by any two time-like sections, not necessarily parallel to one another. The vectors $p_{\text{el}}^\sigma(1)$, $p_{\text{el}}^\sigma(2)$ in 3.9 are now the electromagnetic energy and momentum vectors defined in two different Lorentz frames of reference (frames in which the two sections respectively are sections of constant time). Combined with the (integrated) equation of motion 3.10, this gives

$$p^\sigma(1) = p^\sigma(2),$$

which now expresses, not only the conservation of energy and momentum, but also the fact that the definition of p^σ is independent of the co-ordinate system.

4. FINITE ENERGY AND MOMENTUM—EQUATIONS OF MOTION

It must now be shown that a tensor $K^{\mu\nu\sigma}$ exists. We shall use the following notation:

x^μ are the co-ordinates of a point in space, at which the fields, etc., are being evaluated.

z^μ are the co-ordinates of the point where the world-line of the charge is cut by the *past* branch of the light-cone starting from the point x^μ ; the point z^μ will be called the retarded point.

v^μ is the four-velocity \dot{z}^μ , evaluated at the retarded point.

\dot{v}^μ , \ddot{v}^μ are the derivatives of the velocity with respect to τ , at the retarded point.

s^μ is the null-vector joining the retarded point to the variable point ($s^\mu = x^\mu - z^\mu$).

s is the invariant "retarded distance" $s^\mu v_\mu$.

$\kappa = s^\mu \dot{v}_\mu$.

The inner product of two vectors will be written in bracket notation: $A^\mu B_\mu = (A, B)$.

The actual field will be written as the sum of the retarded field of the charge and the "incoming" external field

$$F = F_{\text{ret}} + F_{\text{in}}. \quad (4.1)$$

The retarded field is given by

$$F_{\text{ret}}^{\mu\nu} = e \left\{ \frac{(1-\kappa)}{s^3} (s^\mu v^\nu - s^\nu v^\mu) + \frac{1}{s^2} (s^\mu \dot{v}^\nu - s^\nu \dot{v}^\mu) \right\}. \quad (4.2)$$

In the neighbourhood of the charge, s tends to zero, and it is convenient to classify the orders of infinity of the various terms with the help of the powers of s ; s , s^μ , κ are $O(s)$, and v^μ , \dot{v}^μ , \ddot{v}^μ are $O(1)$. The retarded field is $O(s^{-2})$.

The energy tensor $T^{\nu\sigma}$ is quadratic in $F_{\mu\nu}$, and can therefore be written

$$T = T_{\text{ret}} + T_{\text{mix}} + T_{\text{in}}, \quad (4.3)$$

where T_{ret} , T_{in} are the tensors formed from F_{ret} , F_{in} according to 3.1, and T_{mix} is the cross-term. T_{ret} is $O(s^{-4})$. T_{mix} and T_{in} are $O(s^{-2})$ and $O(1)$ respectively, and are therefore integrable. Thus $K^{\mu\nu\sigma}$ must be such that the terms of order s^{-4} and s^{-3} in both $T_{\text{ret}}^{\nu\sigma}$ and $\partial K^{\mu\nu\sigma}/\partial x^\mu$ are equal.

$T_{\text{ret}}^{\nu\sigma}$ is given by

$$T_{\text{ret}}^{\nu\sigma} = \frac{e^2}{4\pi} \left\{ -\frac{(1-\kappa)^2 s^\nu s^\sigma}{s^6} + \frac{(1-\kappa)(v^\nu s^\sigma + v^\sigma s^\nu)}{s^5} + \frac{(\dot{v}^\nu s^\sigma + s^\nu \dot{v}^\sigma) - (\dot{v}, \dot{v}) s^\nu s^\sigma - \frac{1}{2} g^{\nu\sigma}}{s^4} \right\}. \quad (4.4)$$

In the appendix it is shown that the divergence of the tensor

$$K^{\mu\nu\sigma} = \frac{e^2}{16\pi} \left\{ \frac{9\kappa(s^\mu v^\nu - s^\nu v^\mu) s^\sigma}{s^5} + \frac{(v^\nu g^{\mu\sigma} - v^\mu g^{\nu\sigma})}{s^3} \right. \\ \left. + \frac{(1+2\kappa)(s^\mu g^{\nu\sigma} - s^\nu g^{\mu\sigma}) + 3(v^\mu s^\nu - v^\nu s^\mu) v^\sigma + 3(\dot{v}^\mu s^\nu - \dot{v}^\nu s^\mu) s^\sigma}{s^4} \right\} \quad (4.5)$$

$$\text{is } \frac{\partial}{\partial x^\mu} K^{\mu\nu\sigma} = \frac{e^2}{4\pi} \left\{ \frac{(8\kappa^2 + 2\kappa - 1) s^\nu s^\sigma}{s^6} + \frac{(\dot{v}^\nu s^\sigma + s^\nu \dot{v}^\sigma) - \frac{1}{2} g^{\nu\sigma}}{s^4} \right. \\ \left. + \frac{(1-\kappa)(s^\nu v^\sigma + v^\nu s^\sigma) - 2(\dot{v}, s) s^\nu s^\sigma}{s^5} \right\}, \quad (4.6)$$

which differs from 4.4 only by terms of order s^{-2} . Defining a $T'_{\text{ret}}{}^{\nu\sigma}$ analogous to $T'^{\nu\sigma}$,

$$T'_{\text{ret}}{}^{\nu\sigma} = T_{\text{ret}}{}^{\nu\sigma} - \frac{\partial}{\partial x^\mu} K^{\mu\nu\sigma}, \quad (4.7)$$

$$\text{this gives } T'_{\text{ret}}{}^{\nu\sigma} = \frac{e^2}{4\pi} \left\{ -\frac{9\kappa^2}{s^6} + \frac{2(\dot{v}, s)}{s^5} - \frac{(\dot{v}, \dot{v})}{s^4} \right\} s^\nu s^\sigma. \quad (4.8)$$

To ensure that the amount of energy radiated away from the charge during its past history be finite, let it be assumed that previous to some remote time it was moving uniformly in a straight line. $K^{\mu\nu\sigma}$ is then of order R^{-3} at infinity, for at sufficiently large distances \dot{v}^μ and κ vanish; equation 3.6 is therefore valid. The expression for the electromagnetic energy is then

$$p_{\text{el}}^\sigma = \int (T'_{\text{ret}}{}^{0\sigma} + T_{\text{mix}}^{0\sigma} + T_{\text{in}}^{0\sigma}) dx^1 dx^2 dx^3. \quad (4.9)$$

This is finite. For a charge in uniform motion, with no external field, it is zero; this is in agreement with the result of Wentzel (1934). The electromagnetic energy is not necessarily positive.

To determine R^σ one must evaluate the integral of $T'^{\nu\sigma}$ along a world-tube of infinitesimal diameter. Only T'_{ret} and T_{mix} will contribute anything, for the integral of T_{in} will tend to zero. The only important terms in T_{mix} are

$$T_{\text{mix}}^{\nu\sigma} \sim \frac{1}{4\pi s^3} \{ (s^\nu v^\tau - s^\tau v^\nu) F_\tau^\sigma + F_\tau^\nu (s^\tau v^\sigma - s^\sigma v^\tau) + \frac{1}{2} (s^\alpha v^\beta - s^\beta v^\alpha) F_{\alpha\beta} g^{\nu\sigma} \}.$$

To evaluate the contribution from an element $\delta\tau$ of the world-line it is simplest to take a co-ordinate system in which the components of v^μ at

the element are $(1, 0, 0, 0)$, and a tube of spherical section. The contribution is then

$$R^\sigma \delta\tau = - \int T'^{\alpha\sigma} d\sigma_\alpha \delta\tau,$$

$$\text{i.e.} \quad R^\sigma = - \int T'^{\alpha\sigma} d\sigma_\alpha, \quad (4.10)$$

where $d\sigma_\alpha$ has the same meaning as on p. 394.

Let x, y, z be the co-ordinates of a point on the sphere and ϵ its radius; then to the required order of approximation

$$s^0 = s = \epsilon, \quad s^1 = x, \quad s^2 = y, \quad s^3 = z \quad (x^2 + y^2 + z^2 = \epsilon^2).$$

The integration can be carried out separately for R^0, R^a ; it is quite straightforward if one remembers that in the particular co-ordinate system chosen $\dot{v}^0 = 0, \quad \ddot{v}^0 = -(\dot{v}, \dot{v})$ (these follow from $(v, \dot{v}) = 0, \quad (v, \ddot{v}) + (\dot{v}, \dot{v}) = 0$). The result is

$$R^0 = 0, \quad R^a = \frac{2}{3}e^2\ddot{v}^a + eF_{\text{in}}^{a0}.$$

In a general co-ordinate system this reads

$$R^\sigma = \frac{2}{3}e^2\{\ddot{v}^\sigma + (\dot{v}, \dot{v})v^\sigma\} + eF_{\text{in}}^{\sigma\nu}v_\nu. \quad (4.11)$$

R^σ satisfies the equation

$$v_\sigma R^\sigma = 0; \quad (4.12)$$

p_{mech}^σ must therefore satisfy the condition

$$v_\sigma \dot{p}_{\text{mech}}^\sigma = 0. \quad (4.13)$$

This is most simply achieved if p_{mech}^σ is a constant multiple of v^σ .

$$p_{\text{mech}}^\sigma = mv^\sigma. \quad (4.14)$$

We have seen that the electromagnetic part of p^σ vanishes for a charge in uniform motion; its total energy and momentum is then therefore mv^σ . The constant m is therefore the actual inertial mass of the charge and its own field.

The equation of motion is

$$m\dot{v}^\sigma - \frac{2}{3}e^2\{\ddot{v}^\sigma + (\dot{v}, \dot{v})v^\sigma\} = eF_{\text{in}}^{\sigma\nu}v_\nu. \quad (4.15)$$

This is the equation given by Dirac (1938, equation 24).

The tensor 4.5 is not the only one that can be built up from the retarded variables to satisfy the requirements. To it could be added

$$\frac{\lambda e^2}{16\pi} \frac{\kappa(s^\mu v^\nu - s^\nu v^\mu) s^\sigma}{s^5},$$

where λ is a numerical factor. The divergence of this is

$$\frac{\lambda e^2}{16\pi} \left\{ \frac{4\kappa^2}{s^6} - \frac{(\ddot{v}, s)}{s^5} \right\} s^\nu s^\sigma,$$

which is also of order s^{-2} . It is easily found that this changes p_{el}^σ by an amount $-\frac{1}{3}\lambda e^2 \dot{v}^\sigma$. The corresponding mechanical vector is $mv^\sigma + \frac{1}{3}\lambda e^2 \dot{v}^\sigma$. The tensor 4.5 was chosen to satisfy 4.12, so that the mechanical part of p^σ could be taken to be mv^σ , without any additional terms.

5. SYMMETRY BETWEEN RETARDED AND ADVANCED VARIABLES

Up to this point the theory has been developed in an unsymmetrical way as regards the advanced and retarded fields. It is possible to formulate it quite symmetrically if one forms a tensor like 4.5 in which the advanced variables occur instead of the retarded. Let us call these two K_{ret} and K_{adv} . In the appendix it is shown that near the world-line $K_{\text{ret}} - K_{\text{adv}}$ is $O(s^{-2})$, and that, further, its integral over a small closed surface tends to zero. It is therefore immaterial which of the two is used in defining p_{el}^σ , and a symmetrical presentation is achieved by using

$$K_{\text{sym}} = \frac{1}{2}(K_{\text{ret}} + K_{\text{adv}}). \quad (5.1)$$

Let the field $F^{\mu\nu}$ be written in the form

$$F = \frac{1}{2}(F_{\text{ret}} + F_{\text{adv}}) + f, \quad (5.2)$$

and let $T_{\text{sym}}^{\nu\sigma}$ denote the energy tensor formed from $\frac{1}{2}(F_{\text{ret}} + F_{\text{adv}})$ according to 3.1, $t^{\nu\sigma}$ that from f , and $t_{\text{mix}}^{\nu\sigma}$ the cross-term. We define $T_{\text{sym}}^{\nu\sigma}$ by

$$T_{\text{sym}}^{\nu\sigma} = T_{\text{sym}}^{\nu\sigma} - \frac{\partial}{\partial x^\mu} K_{\text{sym}}^{\mu\nu\sigma}.$$

In evaluating R^σ by integrating

$$\int (T_{\text{sym}}^{\nu\sigma} + t_{\text{mix}}^{\nu\sigma} + t^{\nu\sigma}) dS_\nu,$$

over a world-tube, it is easily seen from the symmetry that the contribution from $T_{\text{sym}}^{\nu\sigma}$ is zero, and the only contribution comes from $t_{\text{mix}}^{\nu\sigma}$, namely

$$R^\sigma = e f^{\sigma\nu} v_\nu.$$

This gives the symmetrical form of the equation of motion (Dirac, 1938, equation 22),

$$m\dot{v}^\sigma = e f^{\sigma\nu} v_\nu. \quad (5.3)$$

I am greatly indebted to Professor Dirac for letting me use the manuscript of his paper.

APPENDIX

The rules for differentiating the retarded variables are easily obtained by differentiating

$$s^\nu = x^\nu - z^\nu,$$

giving
$$\frac{\partial s^\nu}{\partial x^\mu} = \delta_\mu^\nu - v^\nu \frac{\partial \tau}{\partial x^\mu}.$$

Now s^ν is a null-vector; hence
$$s_\nu \frac{\partial s^\nu}{\partial x^\mu} = 0.$$

I.e.
$$\frac{\partial \tau}{\partial x^\mu} = \frac{s_\mu}{s},$$

$$\frac{\partial s^\nu}{\partial x^\mu} = \delta_\mu^\nu - \frac{v^\nu s_\mu}{s}.$$

The derivatives of v^ν , \dot{v}^ν and s are

$$\frac{\partial v^\nu}{\partial x^\mu} = \frac{\dot{v}^\nu s_\mu}{s}, \quad \frac{\partial \dot{v}^\nu}{\partial x^\mu} = \frac{\ddot{v}^\nu s_\mu}{s},$$

$$\frac{\partial s}{\partial x^\mu} = v_\mu - \frac{(1-\kappa)s_\mu}{s}.$$

With these rules it is easy to derive the following results:

$$\begin{aligned} \frac{\partial}{\partial x^\mu} \left[\frac{\kappa(s^\mu v^\nu - s^\nu v^\mu) s^\sigma}{s^5} \right] &= \frac{4\kappa^2 s^\nu s^\sigma}{s^6} - \frac{(\ddot{v}, s) s^\nu s^\sigma}{s^5}, \\ \frac{\partial}{\partial x^\mu} \left[\frac{s^\mu g^{\nu\sigma} - s^\nu g^{\mu\sigma}}{s^4} \right] &= \frac{-4(1-\kappa)s^\nu s^\sigma}{s^6} + \frac{v^\nu s^\sigma + 4s^\nu v^\sigma}{s^5} - \frac{2g^{\nu\sigma}}{s^4}, \\ \frac{\partial}{\partial x^\mu} \left[\frac{\kappa(s^\mu g^{\nu\sigma} - s^\nu g^{\mu\sigma})}{s^4} \right] &= \frac{-4\kappa(1-\kappa)s^\nu s^\sigma}{s^6} + \frac{\kappa(v^\nu s^\sigma + 4s^\nu v^\sigma) - (\ddot{v}, s) s^\nu s^\sigma}{s^5} \\ &\quad - \frac{\kappa g^{\nu\sigma} + s^\nu \dot{v}^\sigma}{s^4}, \\ \frac{\partial}{\partial x^\mu} \left[\frac{(v^\mu s^\nu - v^\nu s^\mu) v^\sigma}{s^4} \right] &= \frac{-3\kappa s^\nu v^\sigma}{s^5} + \frac{v^\nu v^\sigma + s^\nu \dot{v}^\sigma}{s^4}, \\ \frac{\partial}{\partial x^\mu} \left[\frac{(\dot{v}^\mu s^\nu - \dot{v}^\nu s^\mu) s^\sigma}{s^4} \right] &= \frac{4\kappa(1-\kappa)s^\nu s^\sigma}{s^6} + \frac{(\ddot{v}, s) s^\nu s^\sigma - \kappa(v^\nu s^\sigma + s^\nu v^\sigma)}{s^5} \\ &\quad + \frac{\dot{v}^\nu s^\sigma + s^\nu \dot{v}^\sigma}{s^4}, \\ \frac{\partial}{\partial x^\mu} \left[\frac{v^\nu g^{\mu\sigma} - v^\mu g^{\nu\sigma}}{s^3} \right] &= \frac{3(1-\kappa)v^\nu s^\sigma}{s^5} + \frac{2\kappa g^{\nu\sigma} - 3v^\nu v^\sigma + \dot{v}^\nu s^\sigma}{s^4}. \end{aligned}$$

Multiplying these by 9, 1, 2, 3, 3, 1 respectively, adding and multiplying the sum by $e^2/16\pi$, we find

$$\begin{aligned} \frac{e^2}{16\pi} \frac{\partial}{\partial x^\mu} & \left\{ \frac{9\kappa(s^\mu v^\nu - s^\nu v^\mu) s^\sigma}{s^5} + \frac{(v^\nu g^{\mu\sigma} - v^\mu g^{\nu\sigma})}{s^3} \right. \\ & \left. + \frac{(1+2\kappa)(s^\mu g^{\nu\sigma} - s^\nu g^{\mu\sigma}) + 3(v^\mu s^\nu - v^\nu s^\mu) v^\sigma + 3(\dot{v}^\mu s^\nu - \dot{v}^\nu s^\mu) s^\sigma}{s^4} \right\} \\ = \frac{e^2}{4\pi} & \left\{ \frac{(8\kappa^2 + 2\kappa - 1) s^\nu s^\sigma}{s^6} + \frac{(\dot{v}^\nu s^\sigma + \dot{s}^\nu v^\sigma) - \frac{1}{2} g^{\nu\sigma}}{s^4} \right. \\ & \left. + \frac{(1-\kappa)(s^\nu v^\sigma + v^\nu s^\sigma) - 2(\ddot{v}, s) s^\nu s^\sigma}{s^5} \right\}, \end{aligned}$$

which is the result quoted in 4.6.

To compare K_{ret} and K_{adv} near the world-line let us express them in terms of new variables. Let y^μ be the point on the world-line such that $\gamma^\mu = x^\mu - y^\mu$ is orthogonal to the velocity λ^μ , taken at y^μ . Let $(\gamma, \gamma) = -r^2$, and let τ be the difference in the proper times at z^μ and y^μ . Then to the required order of small quantities

$$s^\mu = \gamma^\mu + \tau \lambda^\mu - \frac{1}{2} \tau^2 \dot{\lambda}^\mu + O(r^3),$$

$$\tau = r \{1 + \frac{1}{2}(\gamma, \dot{\lambda})\} + O(r^3),$$

$$v^\mu = \lambda^\mu - r \dot{\lambda}^\mu + O(r^2),$$

$$\dot{v}^\mu = \dot{\lambda}^\mu + O(r),$$

$$s = r \{1 - \frac{1}{2}(\gamma, \dot{\lambda})\} + O(r^2),$$

$$\kappa = (\gamma, \dot{\lambda}) + O(r^2).$$

Calculating $K_{\text{ret}}^{\mu\nu\sigma}$ to the order r^{-2} we find

$$\begin{aligned} K_{\text{ret}}^{\mu\nu\sigma} = \frac{e^2}{16\pi} & \left\{ \frac{9(\gamma, \dot{\lambda}) \gamma^\mu \lambda^\nu \gamma^\sigma}{r^5} + \frac{\{1 + 4(\gamma, \dot{\lambda})\} \gamma^\mu g^{\nu\sigma} + 3\{1 - (\gamma, \dot{\lambda})\} \lambda^\mu \gamma^\nu \lambda^\sigma + 3\dot{\lambda}^\mu \gamma^\nu \gamma^\sigma}{r^4} \right. \\ & \left. + \frac{3(\gamma, \dot{\lambda}) \lambda^\mu g^{\nu\sigma} + 3\gamma^\mu \lambda^\nu \dot{\lambda}^\sigma + 3\dot{\lambda}^\mu \lambda^\nu \gamma^\sigma}{r^3} + \frac{\dot{\lambda}^\mu g^{\nu\sigma} + 3\dot{\lambda}^\mu \lambda^\nu \lambda^\sigma}{2r^2} \right\} - \dots + O(r^{-1}), \end{aligned}$$

where the minus sign at the end indicates that the similar terms with μ and ν interchanged must be subtracted.

The corresponding result for K_{adv} is obtained by reversing the sign of λ and leaving that of $\dot{\lambda}$ unchanged. The difference of the two is

$$\begin{aligned} K_{\text{ret}}^{\mu\nu\sigma} - K_{\text{adv}}^{\mu\nu\sigma} = \frac{3e^2}{8\pi} & \left\{ \frac{3(\gamma, \dot{\lambda}) \gamma^\mu \lambda^\nu \gamma^\sigma}{r^5} \right. \\ & \left. + \frac{(\gamma, \dot{\lambda}) \lambda^\mu g^{\nu\sigma} + \gamma^\mu \lambda^\nu \dot{\lambda}^\sigma + \dot{\lambda}^\mu \lambda^\nu \gamma^\sigma}{r^3} - \dots + O(r^{-1}), \right\} \end{aligned}$$

which is $O(r^{-2})$.

To calculate $\int (K_{\text{ret}}^{a0\sigma} - K_{\text{adv}}^{a0\sigma}) d\sigma_a$ it is best to take a co-ordinate system in which λ has components (1, 0, 0, 0) and to integrate over a sphere. The components of $d\sigma_a$ are then $\epsilon \gamma^a d\Omega$, where $d\Omega$ is the element of solid angle on the sphere, γ^a the co-ordinates of a point on the sphere, and ϵ the radius. For $\sigma = 0$ the integrand vanishes; for $\sigma = b$ the integral is

$$\int \left(-3 \frac{\lambda^a \gamma^a \gamma^b}{\epsilon^2} + \lambda^b \right) d\Omega + O(\epsilon) = O(\epsilon).$$

Both tensors therefore give the same result for the energy and momentum.

REFERENCES

- Dirac 1938 *Proc. Roy. Soc. A*, **167**, 148.
 Wentzel 1933 *Z. Phys.* **86**, 479, 635.
 — 1934 *Z. Phys.* **87**, 726.

The porous diaphragm method of measuring diffusion velocity, and the velocity of diffusion of potassium chloride in water

BY G. S. HARTLEY AND D. F. RUNNICKLES

University College, London

(Communicated by F. G. Donnan, F.R.S.—Received 12 July 1938)

The velocity of diffusion of substances in dilute solution is a very informative property, but one of which it has been possible in the past to make but little use, on account of the very low accuracy of all but the most extended and laborious experiments, as is well seen in the collection of data for KCl in water from various authors plotted by McBain and Dawson (1935). The chief experimental difficulty has been the elimination of convection which will be produced to some extent by any temperature fluctuations or vibration however small, and particularly by the process of dividing up the diffusing solution when this is necessary for analysis. Convection is least serious when concentrated solutions of a heavy solute are examined, because the large density gradient has a stabilizing influence: it is also less serious the more rapidly the solute diffuses. Unfortunately, however, there is as

yet no satisfactory method of interpreting results in concentrated solutions, and the method of diffusion is most in demand for slow diffusing colloidal solutes.

The technique, introduced by Northrop and Anson (1929), and later applied by several other workers (McBain and Liu 1931; Laing-McBain 1933; McBain and Dawson 1935; Valkó 1935; Cole and Gordon 1936; Mouquin and Cathcart 1935), of allowing the diffusion to take place only within a porous diaphragm, constitutes a very great advance. Convection within the diaphragm is almost completely eliminated and is generally utilized outside the diaphragm to maintain the two outer reservoirs of solution uniform in composition. As a consequence of this uniformity we have an approximately steady state in the diaphragm. The method therefore includes the advantage in mathematical interpretation of the steady state method developed by Clack (1908-24), namely, that the rate of diffusion across the diaphragm is an explicit function of the concentration c and the diffusion coefficient D for each concentration, being equal to

$$k \int_{c'}^c D dc, \quad (1)$$

where c' and c are the concentrations in the reservoirs and k is a constant of the dimensions of length. In Clack's apparatus k is the ratio of the cross-sectional area to length of the diffusion tube. In the diaphragm apparatus the value of k can only be determined by calibration with a solute of known diffusion coefficient.

This function permits exact calculation of the value of D for any particular concentration when measurements have been made over a range of concentrations. In non-steady state diffusion measurements, exact calculation of the diffusion coefficient, when this varies with concentration, is prohibitively difficult, as has been explained by Gordon (1937). The tables of Stefan (1879) and Kawalki (1894), which are usually employed to evaluate the diffusion coefficient from experimental results in the non-steady state, were worked out on the assumption that D was constant throughout the experiment and could not, owing to several complications, be worked out on any other assumption.

The results discussed here were for the most part obtained several years ago by one of us (G. S. H.) in an investigation of the porous diaphragm method. The first experiments, made before the publication of Northrop and Anson's work, compared the rates of diffusion of N/5 hydrochloric acid and N/2 potassium chloride into water through a no. 4 Jena G filter. The ratio of the former to the latter, quite consistently obtained, was about

25 % in excess of that predicted and found by Öholm (1904). The method was therefore temporarily abandoned, but when normal and satisfactory results were reported by other workers we repeated the measurements with other Jena G filters (both 3 and 4 porosity) and ourselves obtained normal results in nearly all cases. One other no. 4 diaphragm, however, behaved anomalously in permitting sodium chloride to diffuse apparently almost as rapidly as the potassium salt: it did not however give constant results for any electrolyte examined. During the work reported in the following paper when many more diaphragms were in use, it was found again that a few had to be discarded owing to failure to give reproducible results. We mention this to give particular emphasis to the necessity in this method of using several units of apparatus.

The apparatus used was of a different type from that of Northrop and Anson and McBain. It seemed probable that, when very dilute solutions were examined, gravity could not be relied on to produce uniformity of the solutions on either side of the filter, right up to the surfaces of the latter.* Accordingly, an apparatus was designed in which the stirring was performed mechanically. This permitted also easy adaptation to a method of analysis by conductivity which could be carried out without interruption of the experiment.

One unit is shown in fig. 1. The filter disk *A* (30 mm. diameter) formed the base of a bottle-shaped vessel of Jena G glass of about 35 c.c. capacity. The disks were obtained from the makers sealed into the middle of long tubes. For use in this type of apparatus these were cut near to the disk and then ground level with it. The bottle was held, by means of the extension *B* of its stopper and a rubber bung *C*, inside an outer vessel of pyrex which it fitted fairly closely. When in use the bottle was completely filled with water or dilute solution and the outer vessel with the more concentrated solution up to the level *D*, above the disk but below the stopper. The apparatus was immersed below the level *E* of water in a thermostat at 18° C. This was slightly below the room temperature so that there was little tendency to condensation in the upper part of the outer vessel. The vessel was supported

* The same question has been raised by Mouquin and Cathcart, who also used a rotational apparatus in consequence. The rotation in their case was about a diameter of the diaphragm as axis. Mehl and Schmidt (1937) appear to have misunderstood the difficulty since they stress that, in their experiments by the stationary method, tests showed the liquid in the reservoirs was uniform. We do not doubt that it will be uniform in the bulk of the liquid, but there must be a more or less stagnant layer immediately adjacent to the diaphragm surface so that the diffusion layer is effectively slightly increased. It seemed to us that the thickness of this stagnant layer is to be expected to increase as the density difference of the solutions is decreased.

by fitting the glass cam *F* in a brass chuck, the upper part resting in a suitable saddle. The axis of the vessel was inclined about 30° to the vertical and the chuck, of which four were housed in one tank, was rotated uniformly by an electric motor and gearing. Rotation of the vessel about the oblique axis caused the solutions to be stirred by two pyrex glass spheres, a heavy one *G* in the bottle and a floating one *H* below the disk.

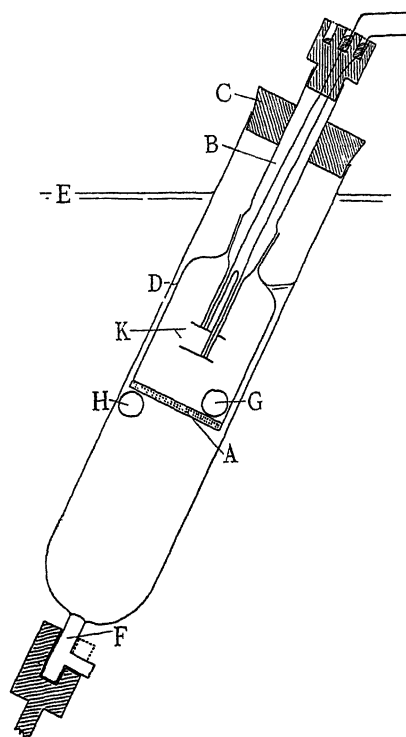


FIG. 1

The stopper of the bottle carried two Arrhenius type platinum electrodes *K*, 1 cm. in diameter, about 1 cm. apart and the lower one about 1.5 cm. above the disk. Since the system was not perfectly symmetrical, the stopper was always turned till the electrodes were in the same position with respect to a mark on the vessel. The resistance was found to vary only slightly with the position of the electrodes so that a very exact method of fixing this position was not necessary. The ebonite contact block *L*, which had two concentric channels cut in it, was fixed to the stopper extension. The leads from the electrodes terminated in these channels, which contained mercury. Stout copper wires, suitably mounted on insulating supports, dipped into the

mercury and permitted connexion to be made between the electrodes and a Wheatstone bridge while the vessel was rotating.

The resistance cell constant of each cell was measured by filling both vessels with standard dilute potassium chloride solutions, using six concentrations ranging from 0.0004 to 0.0014 N. A graph was then constructed for each unit, plotting $\log_{10}(\text{concentration}) + \log_{10}(\text{resistance})$ against $\log_{10}(\text{resistance})$. The first function is of course approximately constant, and a very open scale was used. The slight variation included not only the variation of equivalent conductivity with concentration but also the variation of error in resistance measurement, due to non-ideal leads, resistance-and-capacity-in-series shunts and capacitative coupling between oscillator and bridge. Conversion of resistance measurement to concentration was thus easily made by subtracting $\log_{10}(\text{resistance})$ from the corresponding sum of logs on the graph and taking the antilog. A further simplification was made by considering the residual conductivity of the water used as being in effect due to a small additional concentration of potassium chloride. This procedure was justified since it is only the *differences* between concentrations at different times during each experiment which have to be measured with high accuracy.

The stoppers were used without grease, since the rate of diffusion through the no. 4 diaphragms is seriously reduced by slight contamination. Surface tension was relied on to prevent entry of air through the ground joint, and the slight pressure difference prevented any flow of solution into the bottle. The procedure of an experiment was as follows. First to wash through the diaphragm a large volume of pure deaerated water, and fill the bottle with the dilute solution, also previously deaerated in a filter flask. After insertion of the stopper, the outside of which was dried with filter paper, the bottle was placed for half an hour in a beaker containing some of the solution which was to be used in the outer vessel of the apparatus. During this stage the steady gradient in the diaphragm was approximately established. (It could then be assumed that all the salt diffusing into the bottle during the experiment came from the outer vessel and the decrease in concentration of the latter could then be calculated from the measured capacity of the bottle and the standard volume used in the outer vessel.) The standard volume of concentrated solution was now introduced into the outer vessel, the bottle fitted in, and the whole placed in the thermostat and set in motion.

After allowing at least 2 hr. to elapse, a time found from preliminary measurements to be sufficient for the steady state to be fully established, readings were commenced. These were taken in groups of four with intervals of about 10–20 min. between each reading in a group and of a time between

each group depending on the concentration gradient in use, varying from 2 or 3 hr. when using a N/10 solution in the outer vessel to a day when a N/400 solution was used.

These intervals are considerably less than those used by most workers, but the absence of any disturbance during the experiment enabled very high reproducibility to be obtained. The short intervals permitted a simplification of the method of calculation.

We will denote the concentrations in the bottle by c' and those in the outer vessel by c , v' and v being the corresponding volumes. Numerical suffices to the c 's denote values at corresponding times t , suffix 0 denoting the values at the beginning of the experiment. Then we have

$$v \frac{dc}{dt} = -v' \frac{dc'}{dt}, \quad (2)$$

$$c = c_0 - \frac{v'}{v} (c' - c'_0), \quad (3)$$

$$v' \frac{dc'}{dt} = k \int_{c'}^c D dc. \quad (4)$$

Our first approximation is to assume that the changes in c and c' are too small for there to be a significant change in the mean diffusion coefficient \bar{D} defined by

$$\bar{D}(c - c') = \int_{c'}^c D dc. \quad (5)$$

We may write (4) in the form

$$\frac{dc'}{dt} = K \bar{D}(c - c'), \quad (6)$$

$K = k/v'$ being a characteristic constant of each cell. Integrating (6) with the aid of (2) we obtain

$$K \bar{D} = \frac{1}{(1 + v'/v)(t_2 - t_1)} \log \frac{c_0 + (v'/v)c'_0 - (v'/v + 1)c'_1}{c_0 + (v'/v)c'_0 - (v'/v + 1)c'_2}. \quad (7)$$

In the present calculations we made a further simplification by writing

$$K \bar{D} = \frac{c'_2 - c'_1}{(\bar{c} - \bar{c}')(t_2 - t_1)} = \frac{\delta c'}{\Delta c \delta t}, \quad (8)$$

the average concentrations \bar{c} , \bar{c}' , being $(c_1 + c_2)/2$, $(c'_1 + c'_2)/2$.

By expansion of (7) it can be shown that the first error term in using (8) is

$$\frac{1}{12} \frac{(v'/v + 1)^2 (\delta c')^2}{\Delta c} \frac{c_0 + (v'/v)c_0 + (v'/v + 1)\bar{c}'}{(c_0 + (v'/v)c'_0)^2}, \quad (9)$$

and since, in these experiments, $\delta c'$ amounts to at most 4 % of Δc and v'/v was about 1/3, (8) is certainly sufficiently accurate. In work on electrolytes in dilute solution, where D varies considerably with concentration, error in (7) due to the first approximation (constant \bar{D}) may be more important than the additional error in (8) due to the second. With the short intervals used here, it can be shown that error due to the first approximation is also negligible.

By calculating from the first reading of one group to the last of the next, from the second reading to the next to the last and so on, four completely independent values of $\delta c'/\delta t$ were obtained for each pair of groups and all four had very nearly equal values for the mean concentration difference. A specimen calculation will make this clear and demonstrate the reproducibility obtainable. In the first column of Table I are given the times of the readings in one group and in the second the values of c' corresponding. In the third and fourth columns are given the times and concentrations for the next group in reverse order. Taking the differences between values in each row we obtain the values of $\delta c'$ and δt . In the last column are given the \bar{c}' values for each interval. The times are in minutes and concentrations in g. equiv. per litre at 18° C.

TABLE I

t_1	$c'_1 \times 10^7$	t_2	$c'_2 \times 10^7$	$\delta c'/\delta t \times 10^5$	$\bar{c}' \times 10^6$
139.7	3838	418.5	9164	1910	650
150.2	4036	400.9	8831	1913	643
179.1	4608	386.1	8549	1909	657
197.1	4939	370.7	8257	1911	660

The deviation of the $\delta c'/\delta t$ values from the mean suggests that the fourth figures in the $c' \times 10^7$ values are of doubtful significance even for the determination of concentration increase. The third figures are probably not reliable in the absolute concentration determination, but this is only necessary to a much lower degree of accuracy. Since the \bar{c}' values are nearly the same we take the mean value of $\delta c'/\delta t \times 10^5$, namely 1911, as corresponding to a mean \bar{c}' of 0.00065. In this experiment the initial solution in the bottle was distilled water (corresponding to $c'_0 = 0.00001$), the initial external concentration was 0.10002 and v'/v had the value 0.46, whence $\bar{c} = 0.10002 - 0.46 \times 0.00064 = 0.09973$. Substituting in (8) we obtain

$$K\bar{D}_{0.0006}^{0.0997} = 1929 \times 10^{-4}.$$

The change in \bar{D} during each experiment was less than the experimental error. Calculation for the next group interval in the above experiment yielded, for example,

$$K\bar{D}_{0.0010}^{0.0995} = 1930 \times 10^{-4},$$

a value slightly higher than before, whereas it can be shown from the results for other concentrations that the change in the limiting concentrations must here produce a very slight fall. When lower external concentrations were used, it became possible (and necessary, to secure results of comparable accuracy) to use larger group intervals. A greater relative change in the concentrations resulted, but the drift in \bar{D} as the experiment progressed was still not great enough to be detected with certainty. In tabulating the results we have therefore taken the average value of $K\bar{D}$ for each experiment to correspond to the average concentration limits. These were sensibly the same for all experiments with the same external concentration, and in all cases except that where the external concentration was N/10, the mean internal concentration was the same, 0.0007, the initial internal concentration being adjusted so as to realize this and so keep the cell resistance in the optimum range for measurement.

In experiments with external concentrations N/100 and below it was found necessary to make a correction, amounting to 1.3 % in the case of the N/400 solution, for a slow decrease of resistance found to occur even when the diffusion vessels were filled throughout with the same very dilute solution. This seemed to be due, in part, to slow solution of the glass of the diaphragm, being greater for the finer ones (no. 4), and, in part, to imperfections in the pyrex-platinum seals. To correct for this effect the diffusion vessel was set up with the same very dilute solution throughout and readings taken over a period, not usually more than 24 hr., until the decrease of resistance was constant. The external solution was then changed with the minimum of other disturbance and the diffusion measurement carried out as before. At the end, the external solution was again changed for one made up to the same concentration as that of the internal one and, after a brief period for the excess concentration in the diaphragm to disperse, the "background" decrease of resistance was again measured. The decreases were allowed for by considering them to be, in effect, due to an additional increase of potassium chloride concentration. There was not usually much difference in the rate of increase of this effective concentration at the beginning and end of the experiment: the mean value was subtracted from the total apparent rate of increase during the diffusion.

The results obtained are summarized in Table II. It will be seen from the last two columns that a very high degree of reproducibility is obtained in these measurements. The deviation recorded is that of the average of the four independent results for each group interval from the final average. It is doubtful, however, whether the absolute accuracy is as high as the reproducibility might suggest. The three columns headed $K\bar{D}$ 3 A show the

results obtained separately for three rates of revolution, the subheading figures being the number of seconds per revolution. There is a considerable discrepancy between them at the higher concentrations, higher values being obtained the greater the speed. This indicates that the stirring in some way penetrates the diaphragm itself to a small extent. No such effect could be detected in the no. 4 diaphragm. The ratios of \bar{D} values for the N/40 solution to those for N/10 are 1.029 and 1.030 for the two no. 4 diaphragms and 1.026, 1.027 and 1.025 for the three no. 3 diaphragms at 7 sec. per revolution (cells *B* and *C* were used at this speed only). At 4.1 and 2.9 sec. the ratio for 3 *A* becomes progressively lower, 1.022 and 1.018. It is probable therefore that the difference in the ratio between the no. 3 diaphragm at 7 sec. and the no. 4 is to be attributed to a small residual stirring error in the former.

TABLE II

\bar{c}	\bar{c}'	$K\bar{D} \times 10^4$ (4)		$K\bar{D} \times 10^4$ (3)					No. exps.		No. group intervals		R.M.S. deviation %	
		A	B	A	B	C								
Average	Average			7.0	4.1	2.9			4B	3A	4B	3A	4B	3A
(0.0000)	(0.0007)		(2022)	(2060)										
0.00246	0.00069	—	1986	2028	—	2026	—	—	4	4	12	12	0.3	0.4
0.00495	0.00070	—	1971	2016	2016	2017	—	—	6	6	17	17	0.2	0.3
0.00993	0.00070	—	1948	2000	—	2003	—	—	4	4	10	11	0.2	0.2
0.02489	0.00070	2467	1915	1970	1974	1979	1388	949	9	11	20	25	0.2	0.2
0.04981	0.00074	—	1889	1946	—	1960	—	—	5	4	9	8	0.3	0.3
0.09965	0.00081	2401	1859	1920	1931	1943	1352	926	8	9	15	18	0.2	0.3

In the lower concentrations there is still a discrepancy between the two diaphragms, there being a greater change with concentration in the no. 4, although there is now no appreciable effect of rate of rotation in either. It seemed to us most probable that some kind of surface effect is coming into operation, greater in the finer diaphragm and at the lower concentrations.

We shall select, therefore, to obtain the relationship between diffusion coefficient and concentration, the no. 3 values for the concentrations up to N/100, i.e. as long as there is no appreciable stirring error, and the no. 4 values from N/100 upwards, assuming the surface effect to be then negligible. To permit accurate extrapolation to zero concentration one must know the way in which the curve will approach the D_0 value. It has been shown theoretically (Schreiner 1924; McBain and Liu 1931; Hartley 1931; Onsager and Fuoss 1932; Gordon 1937; not all these authors use the same thermodynamic functions but their equations are equivalent) that the ratio of the diffusion coefficient at a sufficiently low concentration C to that at zero concentration, is given by

$$\frac{D_c}{D_0} = 1 + \frac{d \log f}{d \log C}, \quad (10)$$

where f is the mean activity coefficient, provided that there is no influence of interionic forces on the mobilities. That there is, in diffusion, no first order effect, is generally agreed, but Onsager and Fuoss have shown that there will be a second order effect tending to increase the diffusion velocity. Their derivation of this effect has been called in question (Hermanns 1937; Gordon 1937) and a viscosity correction and volume correction have also been discussed. In concentrations below $N/400$, all these corrections are certainly very small. We therefore felt justified in using the equations of Onsager and Fuoss to provide us with a curve to which to fit our results asymptotically. We have calculated values by the Onsager and Fuoss equation for all our concentrations for comparison. We used for evaluation of the thermodynamic term the equation chosen by Guggenheim (1935) for the osmotic coefficient g , conforming to the Debye-Hückel limiting law and fitting the freezing point data of Adams (1915). The use of the osmotic coefficient is especially convenient for the calculation of \bar{D}_0^c , which is the coefficient most easily compared with our results. We may write, with sufficient accuracy in concentrations below $N/10$,

$$\frac{d \log (fc)}{d \log c} = \frac{d(gc)}{dc}, \quad (11)$$

whence, denoting the Onsager-Fuoss mobility by $\phi(c)$ we have

$$\int_0^c \frac{D}{D_0} dc = \int_0^c [1 - \phi(c)] d(gc), \quad (12)$$

or
$$\frac{\bar{D}_0^c}{D_0} = g - \frac{1}{c} \int_0^c \phi(c) d(gc), \quad (13)$$

the second term in which can be estimated graphically. The values calculated are given in the fourth column of Table III and plotted in fig. 2.

TABLE III

C	\bar{D}_0^c/D_0			Mehl and Schmidt (adjusted)	McBain and others (adjusted)
	$4B$	$3A$	Theor.		
0.0025	0.990	0.984	0.984	1.004	—
0.0050	0.980	0.977	0.978	0.981	—
0.010	0.968	0.968	0.971	0.965	0.972
0.020	—	—	—	0.954	0.954
0.025	0.951	0.953	0.960	—	—
0.050	0.937	0.941	0.951	0.937	—
0.10	0.922	0.909	0.942	0.931	0.924
0.20	—	—	—	0.925	0.917
0.50	—	—	—	0.931	0.903
1.0	—	—	—	0.944	0.949

Our experimental values are proportional to \bar{D}_0^c . A small correction is necessary to bring them proportional to D_0^c (c being here the nearest round concentration to \bar{c}). It was found that the correction from \bar{c} to c was negligible and that \bar{c}' could be taken as 0.0007 in all cases. The correction formula derived from (5) is

$$\bar{D}_0^c = \bar{D}_{0.0007}^c + \frac{0.0007}{c} (\bar{D}_0^{0.0007} - \bar{D}_{0.0007}^c). \quad (14)$$

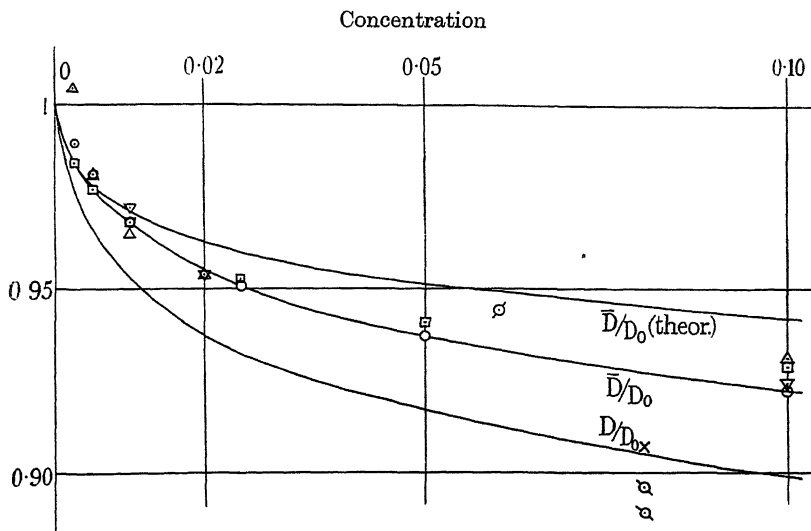


FIG. 2. ○ Present work G4, □ Present work G3, Δ Mehl and Schmidt (adjusted), ▽ McBain and others (adjusted), ⊙ Cohen and Bruins, absolute (\bar{D}/D_0 , probable location), ⊗ Clack, absolute (D/D_0 , corrected for water displacement; to be compared with X).

Sufficiently accurate values of $\bar{D}_0^{0.0007}$ (multiplied by the appropriate K) were obtained by extrapolating $K\bar{D}_0^c$ against \sqrt{c} to $c = 0$ and are given in brackets in Table II. The values obtained for \bar{D}_0^c were divided by constants chosen so that the values for no. 3 and no. 4 diaphragms agreed for $c = 0.01$ and so that the selected series approached tangentially the theoretical \bar{D}_0^c curve. The \bar{D}_0^c/D_0 values obtained by this method of calibration are given in columns 2 and 3 of Table III, the selected series being printed in heavy type.

There is probably very little effect of temperature on \bar{D}_0^c/D_0 in low concentrations, so that our values may be compared with those obtained by McBain and Liu and McBain and Dawson at 25° and by Mehl and Schmidt at 30° C. These workers all used the stationary porous diaphragm

method. Bringing their values arbitrarily into the best average agreement with ours between $c = 0.01$ and $c = 0.10$, we have the figures in columns 5 and 6 of Table III.

There is fairly good agreement as to the variation of diffusion coefficient with concentration in the range from 0.005 to 0.1 N, although the experimental results cannot be strictly comparable, since the values of the other workers quoted have not been corrected as above for the deviation of the concentration limits from the initial values. The agreement found indicates that our doubts about the adequacy of gravitational stirring in the dilute solutions are not justified, or at least that the "stagnant" layers have the same effective thickness at all concentrations.

The actual values recorded by these workers do, however, differ from ours, those of McBain and collaborators being about 3 % higher, while those of Mehl and Schmidt are 8 % lower. This arises from the different methods of calibration used.

Mehl and Schmidt used several methods. Firstly, they extended their measurements to very low concentrations (about N/2000 for KCl) and extrapolated empirically to zero concentration. The limiting slope obtained is some six times greater than the theoretical. The curve strongly resembles those we obtained ourselves in low concentrations before correction was made for the apparent increase of concentration on the dilute side due to solution of glass and other sources of contamination. In the absence of more experimental details it is not possible to attach great weight to these figures.

Their second method was the ingenious one of measuring the electrical conductance through the diaphragm when filled with suitable electrolyte, large platinum electrodes being clamped on either side of the diaphragm for this purpose. This method agreed well with the first, but it is possible that the stagnant layer adjacent to the diaphragm surface in the diffusion measurements is of thickness greater than the distance of separation of the diaphragm surface and the electrodes. Although, as we have seen, there is no change in this thickness with concentration gradient within the limits investigated, the finding of Mouquin and Cathcart that the diffusion through a given diaphragm was some 5 % more rapid when the apparatus was rotated, suggests that the stagnant layer thickness is not negligible.

Their third method was a comparison with Clack's measurement on KCl, but they quote his earlier values (1908) which improved technique (1924) in later work increased by some 10 %. Agreement of our calibration with Clack's more recent measurements is discussed below.

The fourth method was by comparing measurements on 2.0 N NaCl by

the diaphragm method with the results of Öholm at 5° C, corrected for temperature by an equation proposed by Anson.

McBain and collaborators used the value of Cohen and Bruins (1923) for calibration. This was obtained by the "four layer" method, starting with N/10 KCl in the lowest layer and water in the other three. It therefore seems certain, as has been pointed out by Gordon (1937) that the value cannot correspond to $\bar{D}_c^{0.1}$, as is assumed by McBain, but will refer more nearly to a lower concentration range, the mean concentration of all four layers being 0.025 N. The calculation of the diffusion coefficient in the Cohen-Bruins experiment was made by the Stefan-Kawalki tables which were worked out (and could only be worked out) on the assumption that the coefficient is independent of concentration. The effect of this in a case where there is considerable concentration dependence is, as Cohen and Bruins realized, complex. It has the effect, as can be seen from the last column of Table IV, of making the coefficients obtained from layers 2 and 4 (4 is the uppermost) considerably higher than those from layers 1 and 3.

TABLE IV

Vessel	Concentrations for $D =$			Apparent D	Cohen and Bruins	
	0.922	Function	0.951		Concen- tration	D/D_0
1	0.03629	0.03619	0.03581	0.928	0.035	0.945
2	0.02841	0.02831	0.02827	0.944	0.029	0.955
3	0.01992	0.01994	0.02013	0.923	0.021	0.945
4	0.01537	0.01556	0.01579	0.936	0.015	0.951

We have attempted to obtain an approximate allowance for this effect by the following method. Imagine four equal vessels to be connected in series in such a way that diffusion can take place from one to the other under steady state conditions through identical columns and that the diffusion can be interrupted at will. If we allow diffusion to occur for a definite time from vessel 1 to 2 only, then for the same time from 2 to 3 only, then 3 to 4 and back again to 1 to 2 and so on, we obtain finally a distribution not very different from that in the layers in the four layer method. Starting with a concentration of 0.1 in vessel 1 and zero in the others, using a constant value of $0.2 \times \log_e 10$ for the product of time interval and the k of equation (1), and working through eleven cycles of operations, we obtained the figures in columns 2, 3 and 4 of Table IV. Column 2 refers to D (equation (1)) equal to 0.922 throughout ($= \bar{D}_0^{0.1}/D_0$ in our results*). In column 4, D is taken $= 0.951$ throughout. To obtain the values in column 3 we substituted

* The absolute values of D in this calculation are of no consequence, since k is quite arbitrary, and the amount diffusing depends on the product kD .

equation (5) in equation (1), assumed D to be constant only for each separate diffusion period, allotting to it the value of \bar{D}/D_0 from our results for the relevant concentration range. The concentration ranges were taken first to be the means of those in the other calculations, and the cycle of calculations was then repeated to obtain a better approximation.

The concentrations obtained for the four vessels appear sufficiently close to those in the corresponding layers in Cohen and Bruin's experiments to justify our use of them to obtain the effective average concentration. In column 5 we have calculated the *apparent* diffusion coefficients obtained from our hypothetical experiment—i.e. the values which would have been obtained had diffusion proceeded according to the conditions of column 3 but been calculated from the concentrations in column 3 *on the assumption that the diffusion coefficient was constant*. Column 5 is obtained simply by linear interpolation of column 3 between columns 2 and 4. The average value in column 5 is equal to our \bar{D}/D_0 value for 0.6 N. Comparing columns 5 and 7 we see that the discrepancies between the values from the various layers are qualitatively explained. The Cohen-Bruins average value is 1.6 % higher than ours.

It seems to us that convection in the Cohen-Bruins apparatus may have led to too high a value. The great care taken to eliminate vibration and temperature fluctuations is no guarantee that the fundamental difficulty of mixing of the layers during separation was overcome. The effect of the shearing method used for bringing the layers together and separating them was examined only in the initial condition of an experiment when a very great concentration gradient exists in the region of maximum shear. During separation there is only a very small density gradient to prevent non-horizontal flow. The very great reproducibility of the results does not disprove this possibility of mixing, since all the experiments were carried out in exactly the same way and for nearly equal times.

To compare our results with those of Clack, it is most convenient first to convert our \bar{D} 's into D 's, the diffusion coefficients at definite concentrations. From equation (5) we obtain

$$D_c = \bar{D}_0^c + c \frac{d}{dc} (\bar{D}_0^c), \quad (15)$$

which gives a simple graphical method for the conversion. Applied to our results it gives the lower curve on fig. 2. For $c = 0.8$, we find $D/D_0 = 0.905$. This figure must now be corrected to give the diffusion coefficient on the diffusion-relative-to-solvent basis used by Clack, i.e. it must be increased by a fraction cV of its value where V is the partial molal volume of KCl

in dilute solution (0.027 litre). We thus obtain the value 0.907. The figures given by Clack (1924), in what he considers to be his most accurate experiments, work out to $D/D_0 = 0.889$ and 0.896 . The mean is 1.5 % less than our value.

We cannot obtain values for D/D_0 at higher concentrations from our own measurements nor at lower ones from Clack's. If however we take the values of Mehl and Schmidt, adjusted to our method of calibration, differentiated and corrected for solvent transfer, we calculate D/D_0 at $C = 0.10, 0.25, 0.50$ and 1.0 to be $0.923, 0.934, 0.960, 0.997$. The data of McBain and collaborators similarly yield $0.907, 0.923, 0.955, 1.017$.* Clack's mean values give $0.887, 0.872, 0.883, 0.915$. These are considerably lower and might therefore seem to favour Mehl and Schmidt's calibration, which gives values 8 % lower than ours. It must be noticed however that the discrepancy is not constant, that between the adjusted values of Mehl and Schmidt at 30°C and those of Clack at 18.5°C increasing from 3.6 % at $C = 0.10$ to 8.2 % at $C = 1.0$. We suspect that the influence of temperatures on D/D_0 is becoming serious in the concentrated solutions.

It is evident from the foregoing discussion that there is still room for a more thorough and accurate investigation of the diffusion of simple electrolytes, but we think that the values in heavy type in Table III, using the theoretical extrapolation method of calibration and giving results in reasonable agreement with the best work by absolute methods (1.5 % below Cohen and Bruin's value, taken to refer to $\bar{D}_0^{0.06}$, and 1.5 % above Clack's mean value for $C = 0.08$) are at least accurate to 2 %.

THE STIRRING EFFECT IN THE ROTATIONAL APPARATUS

The rotational apparatus described offers great advantages with regard to reproducibility when a conductimetric method of analysis is used, and also the advantage, pointed out by Mouquin and Cathcart, that it is applicable when there is a negligible density gradient in the diffusing solution. It suffers however from the drawback that there is appreciable stirring apparently within the coarser diaphragm.

Preliminary to the work reported in the following paper, we decided to investigate the effect more closely.

* The value of \bar{D} for $C = 0.5$ has been ignored as erroneously low. These workers also determined the differential coefficient directly, finding $D/D_0 = 0.927$ and 0.955 at $C = 0.5$ and 1.0 . In higher concentrations there are much greater discrepancies between the differentiated integral values and the directly determined differential values.

The apparatus was constructed on a sounder mechanical principle. All the gearing was outside the thermostat tank, the diffusion units passing co-axially *through* the large worm-driven final gear wheels. The motive power was a constant-speed A.C. motor and a gear-box enabled the speeds to be varied. We are indebted to Mr G. E. Fensom for the design and construction of the mechanical part of the apparatus. In the diffusion units themselves the outer rubber bung was replaced by a glass stopper, grooved round the middle and greased in the upper part only, and the water level of the thermostat was above this stopper. The electrodes were sealed into Jena N glass. The reproducibility of the measurements was not further improved by these improvements in design. The measurements were made at 25° C.

The first variation investigated was that of the density of the glass stirring spheres, as it was thought that they might press too hard on the diaphragm surfaces. There was no change however in the effect of speed of rotation on the apparent rate of diffusion. A glass pendulum suspended by a glass ball and ring joint from the lower electrode support was also tried, but was found unsatisfactory in the conductimetric analysis method.

TABLE V

Electrolyte and conc.	Revs. per min.	\bar{D} , cm. ² /sec. $\times 10^5$	
		4	3
KCl, 0.05 N	6	1.87	1.88
—	9	1.88	1.91
—	18	1.90	1.99
—	30	1.88	2.36
—	45	1.92	3.76
Mg ₂ SO ₄ , 0.1 N	9	0.65	0.66
„	45	0.64	1.12
Cetyl pyridinium chloride, 0.05 N	9	0.75	0.75
	45	0.74	1.17

The values of the apparent mean diffusion coefficient for three electrolytes are given in Table V. Calibration was effected by the figures of Table III applied to the results for potassium chloride extrapolated to zero speed of rotation. It will be seen that there is hardly any effect of speed of rotation on the results with the no. 4 diaphragm. For the no. 3 (coarser) diaphragm there is a very big effect in the higher speeds. The effect was unusually great in this particular diaphragm. Other no. 3's, used in the work reported in the following paper, showed a much smaller effect. The abnormally great effect, due probably to the presence of a few large pores or cracks, con-

veniently illustrates an important feature of the effect—that it is by no means linear against the speed of rotation, but increases as a considerably higher power than the first.

One factor operating to produce this result is that transfer of solute by convective flow will not, when small, be simply additive to that produced by diffusion. This may be appreciated by considering a simplified diaphragm containing only two equal pores, up one of which and down the other there is a slow flow of solution. As a further simplification we will consider the flow to be uniform, not Newtonian, the velocity being everywhere $= V$. In the steady state of combined diffusion and convection the rates of transfer of solute across normal sections of the pores must be, for each pore, constant throughout its length, the rates being

$$\left. \begin{aligned} A \left(Vc - D \frac{dc}{dx} \right) &= K_1, \\ A \left(-Vc - D \frac{dc}{dx} \right) &= K_2, \end{aligned} \right\} \quad (16)$$

A being the cross-sections of the pores. If the concentrations at the ends of the pores accommodate perfectly to those of the reservoirs, which we will consider to be 0 (for $x = 0$) and c (for $x = X$), we obtain, by integration of (16), assuming D to be constant,

$$\left. \begin{aligned} K_1 &= AVc \frac{1}{1 - e^{VX/D}}, \\ K_2 &= AVc \frac{1}{e^{-VX/D} - 1}. \end{aligned} \right\} \quad (17)$$

Now $K_1 + K_2$ is equal to the total rate of transfer through an area $2A$ and distance X from concentration 0 to c . It is therefore $= -D_{ap} \frac{2Ac}{X}$, where D_{ap} is the apparent diffusion coefficient. Whence

$$\frac{D_{ap}}{D} = \frac{1}{2} \frac{VX}{D} \left[\frac{1}{1 - e^{-VX/D}} - \frac{1}{1 - e^{VX/D}} \right]. \quad (18)$$

For very great values of V , D_{ap} is determined solely by V , but for $VX/D \ll 1$, corresponding to a slight flow superimposed on diffusion, we obtain, by expansion of (18),

$$\frac{D_{ap}}{D} = 1 + \frac{1}{4} \left(\frac{VX}{D} \right)^2 \quad (19)$$

the terms in $(VX/D)^1$ cancelling out. Thus when the error due to convection is small it is proportional to the *square* of the convection velocity. This will not of course apply if the small error is due to excessive convection through a few out of the many capillaries present in the diaphragm. We might expect therefore in practice that the error will at first be proportional to a power of the convection velocity intermediate between 1 and 2.

In the case of the rotational apparatus there are two possible sources of convection—a backward and forward movement through a given part of the diaphragm as it turns from its lowest to its highest position and flow in a constant direction in a given part due to centrifugal force.

The former would depend on the density difference of the upper and lower solutions and would be greater the smaller the speed. There is no evidence of its operation. The latter, whether due to the diaphragm contents having a greater angular velocity than the free liquid (which will be subject to braking action by the glass spheres), or to the density difference of the upper and lower solutions, must be, for given solutions, proportional to the square of the angular velocity.

Combining this conclusion with that above with regard to convection in general, we should expect the error to be proportional to a power of the angular velocity between 2 and 4. This explains the observed result for potassium chloride. It is surprising, however, that the effect is no greater as would be expected from equation (18) for the more slowly diffusing electrolytes.

It is evident from these results and considerations that the rotational apparatus, which has considerable advantages for pure electrolyte work, can be made to give accurate results even with a no. 3 diaphragm, if sufficiently low velocities of revolution are used.

The authors are pleased to express their indebtedness to the Department of Scientific and Industrial Research for a Senior Research Award held by one of them (G. S. H. 1929–32) during his share in the experimental investigation, and to Imperial Chemical Industries for a grant to the other.

SUMMARY

A continuous reading apparatus for measuring the rates of diffusion of pure electrolytes through porous diaphragms is described.

Data obtained for the diffusion of potassium chloride as a function of concentration are presented and the absolute values obtained by extrapolation.

The comparison of the results obtained with the best available data by absolute measurements shows that there is still room for more accurate work in the whole problem, but it is believed that the diffusion coefficients in the range from $N/100$ to $N/10$ are accurate to within 2 %.

REFERENCES

- Adams, L. H. 1915 *J. Amer. Chem. Soc.* **37**, 481-96.
 Clack, B. W. 1908 *Proc. Phys. Soc.* **21**, 374-92.
 — 1911 *Proc. Phys. Soc.* **24**, 40-9.
 — 1914 *Proc. Phys. Soc.* **27**, 56-68.
 — 1921 *Proc. Phys. Soc.* **33**, 259-63.
 — 1924 *Proc. Phys. Soc.* **36**, 313-37.
 Cohen, E. and Bruins, H. R. 1923 *Z. phys. Chem.* **103**, 337-403.
 — — 1924 *Z. phys. Chem.* **113**, 157-9.
 Cole, A. A. and Gordon, A. R. 1936 *J. Phys. Chem.* **40**, 733-7.
 Gordon, A. R. 1937 *J. Chem. Phys.* **5**, 522-6.
 Guggenheim, E. A. 1935 *Phil. Mag.* (7), **19**, 588-643.
 Hartley, G. S. 1931 *Phil. Mag.* (7), **12**, 473-88.
 Hermanns, J. J. 1937 *Rec. Trav. Chim. Pays-Bas*, **56**, 635-57.
 Kawalki, W. 1894 *Ann. Phys. Lpz.*, **52**, 166-90.
 McBain, J. W. and Dawson, C. R. 1935 *Proc. Roy. Soc. A*, **148**, 32-9.
 McBain, J. W. and Liu, T. H. 1931 *J. Amer. Chem. Soc.* **53**, 59-74.
 McBain, M. E. Laing 1933 *J. Amer. Chem. Soc.* **55**, 545-51.
 Mehl, J. W. and Schmidt, G. L. A. 1937 *Univ. Calif. Publ. Physiol.* **8**, 165-87.
 Mouquin, H. and Cathcart, W. H. 1935 *J. Amer. Chem. Soc.* **57**, 1791-4.
 Northrop, J. H. and Anson, M. L. 1929 *J. Gen. Physiol.* **12**, 543-54.
 Öholm, L. W. 1904 *Z. phys. Chem.* **50**, 309-49.
 Onsager, L. and Fuoss, R. M. 1932 *J. Phys. Chem.* **36**, 2689-778.
 Schreiner, E. 1922 *Tidsskr. Kemi Bergv.* **2**, 151-3.
 Stefan, J. 1879 *S.B. Akad. Wiss. Wien*, **79** (2), 161-214.
 Valkó, E. 1935 *Trans. Faraday Soc.* **31**, 230-44.
-

The determination of the size of paraffin-chain salt micelles from diffusion measurements

BY G. S. HARTLEY AND D. F. RUNNICLES

University College, London

(Communicated by F. G. Donnan, F.R.S.—Received 12 July 1938)

In aqueous solutions of paraffin-chain salts—salts, one of the ions in which is a long paraffin-chain with the ionic group at the end—there exists a fairly well defined “critical” concentration (about $N/1000$ when the chain has 16 C atoms), below which the salt behaves as a normal electrolyte. In the neighbourhood of the critical concentration, all physical properties which have been measured with sufficient accuracy suffer a rather abrupt transition (Hartley 1936*a*; Adam 1936; Powney and Addison 1937; Adam and Shute 1938). At concentrations well above the critical, the behaviour is that of a colloidal electrolyte and there is no doubt that the paraffin-chain ions are almost all aggregated into “micelles”.*

The abruptness of the transition is satisfactorily explained, according to the reasoning of Bury and collaborators (Bury and Jones 1927; Bury and Grindley 1929; Bury and Davies 1930; see also Murray and Hartley 1935), if the micelles have a large optimum size, much smaller aggregates not being formed in appreciable amount.

It was not found possible, from measurements of electrical mobility (Hartley, Collie and Samis 1936), to determine the size of the micelles except between very wide limits, on account of the large and uncertain effect of the ionic atmospheres. All we could do was to assume a value for this size based on a picture of the constitution of the micelle and the forces operating to produce it, and to show that this value could reasonably explain the experimental data. The present work was undertaken with a view to obtaining a direct measure of the size of the micelles.

Our picture of the micelle was of a spherical unit with the paraffin chains lying in a chaotic or “liquid” manner in its interior. The evidence for this arrangement (or lack of arrangement) is: (1) that paraffin-chain salt solutions have a very general solvent action for organic substances almost insoluble

* Howell and Robinson (1936) and Howell and Warne (1937) have put forward an alternative theory of the critical concentration. For reasons put forward in another paper (Hartley 1938) we believe it to be quite untenable.

in water (Smith 1932; Hartley 1937; Lawrence 1937; McBain and Woo 1938); (2) that the micelle is produced by the strong mutual attraction of the water molecules "squeezing" the paraffin out of solution; (3) that clear, mobile solutions of paraffin-chain salts show no stream double refraction; (4) that, if the chains were arranged parallel to each other in a pseudo-crystalline unit, the existence of an optimum size of micelle, *required by the experimental evidence*, would be unintelligible.

This last point emerges clearly from the calculations of Meyer and van der Wyk (1937) who assumed the parallel chain arrangement (Thiessen and Szychalski 1931*) as a basis for a statistical treatment of micelle formation and concluded (p. 1324): "Nous ne trouvons à cette hypothèse (optimum size of micelle) aucune base plausible: le fait qu'à l'aide de cette supposition, il est possible d'obtenir des résultats qui concordent assez bien avec l'expérience, semble devoir être considéré comme purement fortuit."

It appears necessary, in view of this misunderstanding, to repeat the explanation (Hartley 1936a, p. 47) of the difference from the point of view of statistical mechanics, between the "liquid" and "crystalline" micelles. Meyer and van der Wyk (1937) make the doubtless correct assumption that, when another paraffin-chain ion attaches itself to an already existing group of ν ions arranged parallel to each other, the loss of interfacial energy is almost independent of the value of ν . The paraffin chain is considered, as a simplifying approximation, as a small cylinder of paraffin, after the manner of Langmuir and Harkins. The *a priori* probability of the attachment will increase proportionally to the area available for attachment. This leads to a distribution of frequencies of occurrence of micelles of different sizes which may have a flat maximum at a certain size, but in which the average size (as well as this optimum size) *increases continuously as the concentration is increased*. If we compare now our fully developed spherical "liquid" micelle with a "crystalline" micelle of the same content, we see that its *a priori* probability is much greater, because, although we impose about equal restrictions on the *position* of the ionic end groups of the constituent paraffin chain ions in the two cases, we impose a much more severe restriction on the *orientation* of the chains in the latter case. The loss of free surface

* The evidence for a crystalline structure of the micelles is based on X-ray measurements. It refers however to soaps in what McBain distinguishes as the "curd" state. The micelle that we are discussing is that found in the clear mobile solution and it appears that Meyer and van der Wyk must have these solutions in mind because they utilize freezing point data by McBain and Betz on some sulphonic acids which are in clear solution at the freezing point. Many paraffin-chain salts do not appear in the curd form at all, coming out of the clear solution on cooling as ordinary macroscopic crystals.

energy is also slightly greater in the formation of the liquid micelle, since its surface, for a given content, is smaller and more nearly entirely ionic. The much smaller groups of ions are, however, either in the same condition as before, having the same probability, or they are themselves chaotic, in which case the chains will touch in one or two places only, and their free surface energy be therefore much increased.

When the large micelle is "liquid", we should expect therefore a much greater frequency for its occurrence relative to that of the smaller micelles than is allowed for in the calculations of Meyer and van der Wyk. This does not mean however that we should expect only an increase in the average size at all concentrations. We have referred to the "fully developed" micelle, by which we mean a spherical micelle as large as is consistent with the ends of the chains being able occasionally to reach the centre without the ionic groups being drawn out of the water into the non-polar interior of the micelle. The radius of the interior will therefore be of the order of the length of a fully extended chain. One or two chains only need at any time be in this extended and radial condition, the remainder filling up the available volume as would the molecules of a liquid paraffin in a very small droplet. An extension of the micelle beyond this size would require, either that some of the ionic groups become submerged in the paraffin interior, against which very strong forces will operate, or that the sphere be extended into an ellipsoid, which will decrease the *a priori* probability of the arrangement, because the permissible orientation of the chains will be more restricted. As the micelle size increases beyond that of the fully developed spherical micelle, the frequency distribution will again approach that derived by Meyer and van der Wyk.

We should expect therefore that there will be a fairly well defined optimum size for the micelles *almost independent of concentration*, and that this size will correspond to a sphere of radius of the order of the length of a fully extended chain.

THE DIFFUSION OF COLLOIDAL ELECTROLYTES

In order to obtain a direct estimate of the size of charged particles from diffusion measurements, it is necessary that these be carried out in a "swamping" excess of a simple electrolyte (Hartley and Robinson 1931; Bruins 1931, 1932). If a solution of the colloidal electrolyte alone is investigated, the mobile "gegenions" compel diffusion to occur at a rate comparable with that of an ordinary electrolyte. The comparison of magnesium sulphate and cetyl pyridinium chloride in Table V of the previous paper provides

a good example of this. This high diffusion coefficient is very insensitive to changes in aggregation of the colloidal ion and is very sensitive to the presence of electrolyte impurities. Its relationship to the mobilities includes the factor $1 + \frac{d \log f}{d \log c}$, where f is the activity coefficient, so that a knowledge of the thermodynamic properties of the solution is necessary before even a limited amount of information can be obtained. In the presence of a swamping excess of simple electrolyte, however, the diffusion potential is eliminated, so that the colloidal ions diffuse independently, the thermodynamic correction disappears (f being constant throughout the system), and influence of electrolyte impurities is negligible. Of the two mobility corrections of the Onsager theory the "endosmotic" one is completely absent, but it is possible that a small relaxation term remains (Hartley 1935), the magnitude of which is a question which still awaits theoretical examination.

A considerable amount of work has been done on the diffusion of dyes in solution, but colloidal electrolytes of the paraffin-chain salt type appear to have received little attention. McBain and Liu (1931) examined potassium laurate and Laing-McBain (1933) sodium oleate, by the porous diaphragm method. They used, however, no additional electrolyte, except a trace of alkali to suppress hydrolysis. The results, though of interest in confirming the high value to be expected for a colloidal electrolyte alone, are consequently of little value for estimating the size of the colloidal aggregate. Jander and Weitendorf (1934) studied the diffusion of the sodium salts of valeric, capric and lauric acids in excess of sodium hydroxide or chloride. All of these, even the valerate, diffused more slowly the higher the concentration of fatty ion, but it appears from the figures that the results in the concentrations below 0.02 N must be erroneously high, as they would correspond to electrical mobilities of the ions considerably greater than is to be expected (for example, to 23 ohms⁻¹ cm.² (g. equiv.)⁻¹ at 10° C for C₉H₁₉COO⁻ and 28 for C₄H₉COO⁻, which would correspond to 28 and 35 at 18° C, while Kohlrausch and Holborn (1916) give only 26 at this temperature for the latter). It is probable that the error was due to convection. The free liquid, four-layer technique was employed and, since the soaps make very little change in the density of aqueous solutions, and the diffusing salt must have some effect on the distribution of the "swamping" salt, it is very likely that the density gradient was not sufficient to stabilize the solution (compare Freundlich and Krüger 1935). In the case of the caprate, the diffusion coefficient, after remaining approximately constant between 0.03 and 0.05 N, suffered a second fall in higher concentrations, the value in

0.15 N solution corresponding to a radius of the particles, according to the Stokes-Einstein equation, of about 14 Å. There was no evidence however that a constant value had been reached. In the case of the laurate, where the micelles would be expected to form in much lower concentration, only results in concentrations below 0.01 N are given, presumably owing to limited solubility (the measurements were made at 10° C).

DETERMINATION OF THE PARAFFIN-CHAIN RADICAL

Since the amount of the diffusing paraffin-chain ion must be determined in the presence of excess of other ions, the distribution of which will also be slightly modified by the diffusion process, non-specific methods of concentration determination, such as conductivity and optical refraction, cannot be employed. A method first developed in collaboration with N. K. Adam was applied. This depends upon the fact that the colour of sulphonphthalein indicators is not only displaced to the alkaline side by paraffin-chain cations but is also qualitatively changed (Hartley 1934). The second change only is utilized. Bromphenol blue was found to be the most suitable indicator. Even very low concentrations (less than 10^{-3} N, and therefore below the normal critical concentration for micelles) of the cetyl pyridinium ion change the normal alkaline purple to a pure blue, the transmission in the red end of the spectrum being displaced or eliminated. The colour change is used as an end-point for the direct titration of paraffin-chain anion against paraffin-chain cation. These form an insoluble salt. When the paraffin-chain cation is removed in this manner, the alkaline purple colour is restored. Both the amount and coagulation of the precipitate are at a maximum at the equivalent end-point of the titration, and are a useful indication of the approach of the colour change. The accuracy of the titration varies with the nature of the paraffin-chain ion. The soaps, for example, give very indefinite and non-stoichiometric end-points, but N/1000 cetyl pyridinium or trimethylammonium ions can be titrated against N/1000 cetyl sulphate, or cetane sulphonate, to within 1 %. At lower concentrations the accuracy is somewhat less. The titrations are not seriously affected by the presence of electrolytes up to a concentration of N/2; above this there is a distinct lag in the restoration of the purple colour and the end-point becomes less sharp.

A stock solution of cetyl pyridinium chloride, standardized electrometrically against silver nitrate, was diluted to N/1000. A solution of cetane sulphonate, remaining clear at room temperature, was prepared from a standardized solution of cetane sulphonic acid by the addition of technical

triethanolamine in slight excess. These two solutions were used as standards for subsequent titrations. The titrations, wherever possible, were performed at room temperature; the alkali cetane sulphonates and cetyl sulphate, by reason of their low solubility, were titrated at 60°. In this case the accuracy of the titration was slightly lower.

Determinations were made on 2 c.c. samples in 15 c.c. test tubes, the opposite solution being run in from a microburette, graduated in hundredths of a c.c. 0.1 c.c. of M/5000 bromphenol blue, made slightly alkaline with ammonia, was used in each titration. The titration tube was placed in a comparator between tubes containing mixtures 3 % on either side of the end-point. The solution was titrated to a colour mid-way between the purple and blue colours shown by the outer tubes. Facilities were provided in the comparator for viewing the solutions both by transmitted and scattered light. The presence of the flocculent precipitate made this desirable. Ordinary electric light is preferable to daylight. The indicator showed appreciable red fluorescence at the end-point.

EXPERIMENTAL

For use with paraffin-chain salts, the type of diffusion cell previously described was modified by having provision for the removal of a portion of the solution for external analysis without interrupting the progress of the diffusion. The cell is shown in fig. 1. The porous diaphragm divides the cell into two compartments. The lower compartment contains the paraffin-chain salt solution and is closed at the bottom by a glass stopper. As the strong emulsifying action of the paraffin-chain salt solution on grease precludes its use in ground glass joints, the outside of the stopper is continued in the form of a tube, which provides an airlock when the cell is immersed in the thermostat. Air is prevented from entering the lower compartment by the surface tension acting at the joint, while leakage of solution past the joint is prevented by the hydrostatic pressure of the thermostat liquid.

Eight cells of this type were employed. They were arranged in two thermostats, four cells in each, and were rotated from a common driving mechanism. Further extension of this mechanism was impracticable, and as it was necessary to increase the number of diffusion cells, in order to obtain sufficient measurements in the time available, a number of stationary diffusion cells were constructed. The work discussed in the previous paper had shown that the use of gravitational stirring was satisfactory. The effect of paraffin-chain salts on the density of the swamping electrolyte is very

small and therefore the actual density gradient may be determined by the redistribution of the latter during the diffusion. In order to ensure that there was sufficient density difference, a deliberate slight inequality of the swamping electrolyte concentration was used.

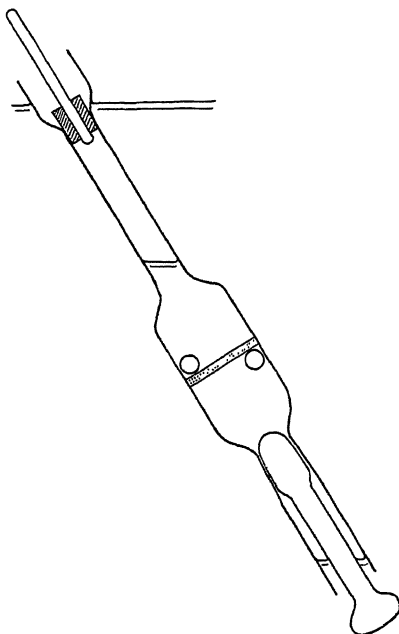


FIG. 1

Twelve stationary type diffusion cells were constructed, and arranged in an air thermostat. The thermostat consists of four long chambers of square section joined together to form a rectangle. Air is circulated round the rectangle by a fan in one of the vertical chambers, the other containing the thermoregulator. The diffusion cells are placed in the horizontal chambers, six in each. A section through one of the horizontal chambers and a cell is shown in fig. 2. Glass doors are provided to enable the cells to be placed in position. A double *V* brass rest (shown in perspective in inset) is fixed to a rigid portion of the thermostat through a universal clamp *D* and the diffusion cell is held in position in it by strong elastic bands. After preliminary levelling of the porous diaphragm with a plumb-bob attachment, the universal clamp is locked, and the diffusion cell can then be removed for cleaning and filling and replaced in exactly the same position. The compartment *A* contains the concentrated solution and is closed at the top by a glass stopper (ungreased) and at the bottom by the

porous diaphragm. *B* is a lipless beaker, the top of which is flattened out in the form of a flange which presses tightly against the rubber bung *C* to form an almost airtight joint. The beaker contains the dilute paraffin-chain salt solution. Removal of this beaker and replacement by one containing fresh solvent is effected from underneath the horizontal chamber as shown in the figure; in this way undue disturbance of the temperature of the thermostat is avoided.

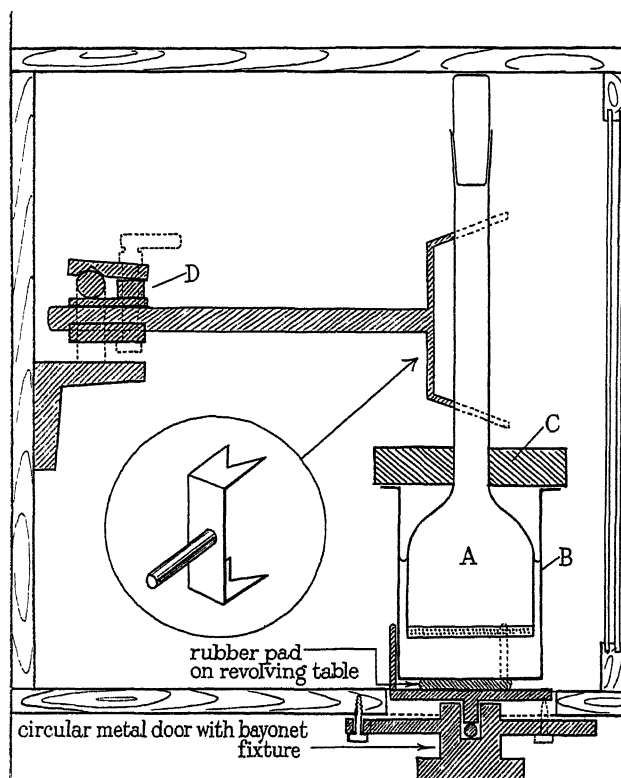


FIG. 2

The diffusion cell constants were determined with 0.05 N potassium chloride, the sample of dilute solution being transferred to a conductivity cell and the salt concentration determined conductimetrically. The lower accuracy of the method of analysis of the paraffin-chain salt solution at low concentrations made it necessary to allow a relatively greater amount of substance to diffuse through the diaphragm between each measurement, with a resultant greater change in the concentration of the concentrated solution. The diffusion coefficients were calculated by the same method as

in the previous paper. Despite the much longer intervals the diffusion coefficients calculated by the approximate equation (8) were not appreciably different from those calculated by the exact equation (7). The more convenient equation (8) was therefore used. As in the measurements described in the previous paper, several determinations of the diffusion coefficient were made with each filling of the diffusion cell, successive samples of the dilute solution being removed after sufficient paraffin-chain salt had diffused through the porous diaphragm. After removal of a sample of dilute solution, it should, ideally, be replaced by one of identical composition. No determinable error, however, resulted from the more convenient procedure of replacing the sample by an equal volume of pure swamping electrolyte, provided that this was allowed for in the calculation of $\delta c'$ and Δc (equation (8)).

It was necessary to make several preliminary measurements to establish that the porous diaphragm method, in the stationary and rotational forms, was satisfactory in its application to the diffusion measurements with paraffin-chain salt solutions. These measurements were made with cetyl pyridinium chloride, prepared as described by Hartley (1936*b*). Sodium acetate was mostly used as the swamping electrolyte; this had the advantage that high concentrations of electrolyte could be used without precipitation of the cetyl pyridinium salt.

TABLE I

	Cetyl pyridinium conc.	Salt and conc.	Type of cell*	Temp.	$\bar{D} \times 10^7$	
					G3	G4
A	0.05	NaCl 0.5	R	30	9.6	9.5
					9.7	9.6
					9.8	9.3
					—	9.6
B	0.01	NaAc 0.1	S	25	9.5	9.2
			R	25	9.2	9.2
C	0.02	NaAc 0.4	R	25	8.8	9.0
	0.02	NaAc 0.4–0.3	R	25	—	8.7
D	0.05	NaAc 1.0	S	30	9.1	9.1
			S†		10.9	9.2

* R=rotational, S=stationary.

† Diaphragm deliberately inclined 8° to horizontal.

The total time of the diffusion measurements was usually from three to four weeks, enabling several determinations to be made. A series of results in two typical experiments is shown in Table I (A). It is seen that within the experimental error, there is no difference between the diaphragms of

different porosity, indicating absence of effects due to absorption in the diaphragm. All the measurements recorded in Table II were made at least in duplicate, using the diaphragms of different porosity, designated G 3 and G 4 by the makers, who give the mean pore diameters as 18 and 4μ respectively.

In Table I (B), the agreement between results with the stationary and rotational apparatus is demonstrated. The figures recorded here, and subsequently, are the means, usually of three determinations, for each experiment. In some cases the means include the results of more than one experiment.

As pointed out above, it was necessary, in the stationary apparatus, to produce adequate density differences by a deliberate inequality of swamping electrolyte concentration. This will in general set up a small potential gradient. The results in Table I (C) show that with a 25 % difference in sodium acetate concentration the diffusion is hardly appreciably decreased. Except in test experiments, a 5 % difference only was used. The error from this source is therefore negligible.

In Table I (D) is shown the effect of deliberately tilting the diaphragm in the stationary apparatus. 8° of tilt produced a 15 % error with the G 3 diaphragm and none with the G 4. This experiment was made in the highest concentration of salt, and a 10 % inequality was used. Since the convection error is (see previous paper) more nearly proportional to the square than to the first power of the angle and of the density difference, and since the diaphragm is levelled to within 1° in the ordinary measurements, the error from this source is also negligible.

SIZE OF CETYL PYRIDINIUM MICELLES

It was seen in the previous section that there was good agreement between the diaphragms of different porosity for cetyl pyridinium chloride with sodium chloride or sodium acetate as swamping electrolyte. The agreement was not disturbed when sodium sulphate or potassium oxalate was used. Moreover, the stationary and rotational methods gave the same results. There can be no doubt therefore that we are measuring the true diffusion coefficient.

Experiments with different gegenions were undertaken largely because a parallel transport number investigation (Samis and Hartley 1938) indicated that in solutions of the paraffin-chain salt alone there were significant differences. Since it soon appeared that the differences, in excess of simple

salt, were small, it was not considered necessary to prepare the appropriate cetyl pyridinium salt in each case, the simple salts being added to cetyl pyridinium chloride. In the case of the sulphate, three measurements (marked † in Table II) were made with no chloride ion present and indeed no change was observed.

It is necessary for the application of the Stokes-Einstein equation that there should be sufficient concentration of simple electrolyte to eliminate the effect of diffusion potential. The concentration necessary might be determined experimentally by increasing the concentration until there is no further fall of diffusion coefficient. Unfortunately, however, there are other effects—that of viscosity, the “atmosphere relaxation” effect and, in some colloidal electrolytes (Valkó 1936), changes of aggregation—which produce a slow steady fall of diffusion coefficient.

Some colloidal electrolytes make a considerable structural contribution to the viscosity and what effect, if any, this has on the diffusion velocity is doubtful. Fortunately, in the present case, the paraffin-chain salt, in the concentrations used, makes very little difference to the viscosity of the simple salt solution to which it is added. The variation of diffusion coefficient with concentration of paraffin-chain salt at a constant concentration of simple salt will therefore be due only to potential and aggregation effects. It will be seen from Table II that there is no such variation outside experimental error. The ratio of simple to colloidal salt is therefore adequate to eliminate the diffusion potential. We have marked with an asterisk the only values in Table II obtained from experiments where this ratio was less than 10 : 1. It was, in these, 5 : 1 and the diffusion coefficient is but little increased.

It must also be concluded that there is no appreciable change of aggregation. We know that, at sufficiently low concentration, the single ions must preponderate in the equilibrium with the micelles. The single ions will have a diffusion coefficient at 25° of about 56×10^{-7} (calculated from the electrical mobility), i.e. about six times that of the micelles. It is obvious, therefore, from the figures in the table, that, even at a concentration of 0.002 N cetyl pyridinium (with 0.04 N sodium acetate), there must be only a very small fraction of the paraffin-chain ions present in the simple form. In pure solutions of cetyl pyridinium chloride, conductivity and transport number (Hartley, Collie and Samis 1936) measurements show that there is no appreciable aggregation at concentrations below 0.0009 N. These conclusions are not in contradiction because the addition of simple electrolyte is known to decrease considerably the critical concentration for micelles (Hartley 1936*a*, p. 37, 1936*b*, 1938*b*).

TABLE II

Cetyl pyridinium conc.	Salt and conc.	Temp.	$\bar{D} \times 10^7$	Radius
Na acetate				
0.002	0.04	25	9.8	24.2
0.005	1.0	25	7.1	24.5
0.01	0.1	25	9.4	24.9
0.01	0.2	25	9.2	24.5
0.01	1.0	25	7.4	23.7
0.02	0.4	25	8.9	23.8
0.02	0.7	25	8.0	24.1
0.02	1.0	25	7.3	24.0
0.05	1.0	25	7.2	24.3
Na chloride				
0.005	0.1	25	8.5	28.4
0.01	0.5	30	9.4	27.9
0.02	0.1	25	9.4*	25.5
0.02	0.2	60	19.4	25.9
0.02	0.4	30	9.9	26.8
0.02	0.5	30	9.4	27.9
0.02	1.0	30	9.0	28.1
0.05	0.5	30	9.6	27.4
Na sulphate				
0.005	0.1	25	8.6	27.9
0.005	0.4	25	8.3†	27.0
0.005	1.0	25	7.7	25.7
0.01	0.2	25	8.6	27.3
0.01	0.4	25	8.3†	27.0
0.01	1.0	25	7.5	26.0
0.02	0.1	25	8.7*	27.5
0.02	0.4	25	8.4	26.8
0.02	0.4	25	8.6†	26.2
0.02	1.0	25	8.0	25.0
K oxalate				
0.02	0.4	30	10.1	26.1
0.02	1.0	30	9.5	25.9
0.05	1.0	30	9.8	25.1

As there is no change in diffusion coefficient with concentration of diffusing electrolyte at any one concentration of swamping electrolyte, the mean diffusion coefficients (\bar{D} of the previous paper) require no correction (to D) before the Stokes-Einstein equation is applied. We have substituted in the Stokes-Einstein equation the viscosity of the simple salt solution used as solvent. The radii thus obtained are recorded in column 6. This allowance for

the effect of the simple salt on viscosity is in the present case more important than in most of the work on dyes, where much lower concentrations have been used. It cannot presumably be exact, if only because the ionic distribution around the micelles means that the solvent is in effect not homogeneous, but it produces so great a unification of the results that it is probably approximately correct. [Compare for example the diffusion coefficient and radius values for 0.02 N cetyl pyridinium and various concentrations of sodium acetate.]

The average deviation of the individual diffusion values from the mean was in general about 2 %, but somewhat greater when the cetyl pyridinium concentration was less than 0.01 N. When sulphate was used as swamping electrolyte the deviation was slightly greater than with the other salts. The method of calibration discussed in the last paper was used. It must be borne in mind therefore that the values of the micelle radius may have an absolute error somewhat greater than that indicated by the internal reproducibility.

It may lastly be noticed that, on increasing the concentration of any one of the swamping electrolytes, there is very little change in the apparent radius, if indeed any, outside the experimental error. We can therefore conclude that the atmosphere relaxation effect, which would operate to reduce the diffusion coefficient, and hence increase the apparent radius, has a negligible effect. Indeed in the case of the sulphate the average apparent radius in 1 N electrolyte is somewhat less than in the lower concentrations. It is extremely improbable that there is a real decrease in the micellar radius, and the observed decrease, which is only just outside the experimental error, is probably well within the error due to substitution of the bulk viscosity in the Stoke-Einstein equation. On account of the ionic distribution about the micelle, this substitution cannot, as pointed out above, be exact, and will be less exact where the gegenions are divalent.

Over a 24-fold range of cetyl pyridinium concentration, from N/500 to N/20, and with added electrolyte ranging from N/25 to N, there is thus no appreciable change in the average size of the micelle. This is exactly what is to be expected from the other evidence on the constitution of these solutions and the arguments outlined in the introduction, the factors involved in micelle formation leading to a well-marked preference for micelles of a very limited range of size. These arguments predict also approximately the size preferred. We should expect the micelle to be spherical having a chaotic paraffin interior of radius equal to the length of a fully extended chain. This length, from the terminal C atom to the ionic group, is about 20 Å. Assuming the density of the interior to be that of liquid hexadecane

(0.77) we calculate that this interior contains about 70 paraffin chains. Allowance must be made now on the surface of this spherical paraffin interior for the ionic pyridinium groups, attached gegenions and the hydration layer. For several reasons this can be made only approximately. If, for instance, the pyridinium is imagined to be in effect spread evenly on the surface, it would extend the radius by only 0.5 Å. This is obviously absurd. The 70 pyridine nuclei must project considerably further than this, and the extension will be increased by the hydration layer, and in many cases, by attached gegenions. Though this extension exists only at some 70 discrete places, its effect on the frictional resistance of the aggregate must be much greater than if the corresponding volume were uniformly spread over the paraffin interior. An effective extension of some 3 or 4 Å is probable.

We must also remember that the paraffin interior of the micelle is "filled" only in the same sense as is a bottle of hexadecane filled by hexadecane molecules. The average distance between neighbouring C atoms of different molecules in hexadecane must be about 5 Å. A mathematic cavity of this diameter in the interior of the micelle may thus not be a cavity in the physical sense. Even if the radius of the interior be taken as great as 23 Å, the thermal perturbations of the outer interface could easily permit sufficiently frequent access of terminal C atoms to the centre. An effective overall radius between 25 and 27 Å, as is found experimentally, is thus not in excess of that required by the constitution of the spherical "liquid" micelle.

The experiments reveal however a definite though small influence of the nature of the gegenions. Thus the average values for acetate oxalate, sulphate and chloride are 24.2, 25.7, 26.6 and 27.2 Å respectively. The radius of the micelle is very little less at 60° than at 25°. It will be evident from the foregoing discussion that the exact constitution of the outside of the micelle will have an appreciable influence on the effective radius, not so much because the different gegenions attached will occupy different volumes as because the differences in the closeness of their attachment will permit variations of the depth to which they, and hence the terminal C atoms of the chain, can penetrate.

Diffusion of anionic micelles

Choline cetane sulphonate, with choline chloride as swamping electrolyte, was the first system examined. It provides a clear solution at room temperatures of a salt with an anion of the same chain length as that of the cation previously examined. This solution was prepared from silver cetane sulphonate by grinding in a mortar with concentrated choline chloride

solution, the requisite excess of choline chloride being added at this stage. The silver chloride was filtered off and the filtrate diluted to the required concentration.

TABLE III. DIFFUSION OF CHOLINE CETANE SULPHONATE AT 25° C

Sulphonate conc.	Chloride conc.	$\bar{D} \times 10^7$		Radius (A)
		G 3	G 4	
0.02	0.4	6.9	—	31
0.05	1.0	6.0	6.5	31

The results, at 25°, are given in Table III. They indicate a micelle size considerably greater than that in the cetyl pyridinium salts. Although there is fair agreement between the results for the diaphragms of different porosities in the case where there are parallel experiments (in the other case the G 4 apparatus in use cracked), a fine precipitate formed during the experiment, due presumably to alkali from the glass, and it seemed probable that this precipitate might interfere with the diffusion, particularly through the finer diaphragms and in long experiments (the second experiment recorded was of shorter duration than usual). It was decided therefore to conduct further experiments at 60° C.

Potassium cetane sulphonate, swamped with potassium chloride, was used, and the measurements carried out in the rotational apparatus in an oil-covered water thermostat. Even in this case a precipitate still formed during the experiments, appearing first on the "dilute" side of the diaphragm. Since this may have been due to abstraction of some alkaline earth metal from the glass, further measurements were made on sodium cetyl sulphate (the alkaline earth salts of which are soluble) swamped with sodium chloride (with the addition of N/400 sodium hydroxide to retard hydrolysis of the ester): no precipitate formed in these solutions.

The individual diffusion values are collected in Table IV, the figures being given in each experiment in the order determined.

There is a very considerable discrepancy between the different diaphragms, even for the sulphate ester, where no precipitate was formed. The values given by the coarser diaphragm are consistently higher than those from the finer, and are generally higher at the commencement of each experiment. This suggests that the real diffusion coefficients will be higher still. The maximum value (of 20×10^{-7}) obtained for the sulphate ester corresponds to a radius of 25 A, about the same as that obtained for the cetyl pyridinium micelle. There is thus no real evidence that the anionic micelle is much larger than the cationic, but the serious systematic dis-

turbance of the diffusion through the diaphragm makes it impossible to arrive at a more definite conclusion.

TABLE IV. DIFFUSION OF ALKALI SALTS AT 60° C

Paraffin-chain conc.	Salt conc.	$\bar{D} \times 10^7$	
		G 3	G 4
Sulphonate 0.01	0.1	12.6	10.5
		12.6	11.0
		11.9	10.1
		11.0	—
0.0067	0.1	10.7	9.5
		11.5	—
0.005	0.1	13.1	9.3
		11.7	6.3
		—	5.1
Sulphate 0.01	0.1	20.4	17.0
		17.7	12.6
		16.4	10.9

A possible source of this disturbance was the existence of a structure which is made evident at higher salt concentrations by the elasticity of the solution in bulk. We decided therefore to attempt some measurements at concentrations where this elasticity is very pronounced. The elasticity rendered it impracticable to use a lighter-than-water glass sphere for stirring the concentrated solution in the diffusion cell. It was necessary to construct a new type of cell. This was similar to the rotational diffusion cell described in the previous paper, but the concentrated solution was placed in the inner vessel and the stirring effected by means of a heavy mercury-filled glass pendulum attached by a ball and ring joint to the glass stopper. Stirring of the dilute solution (which is quite mobile) contained in the outer vessel was effected by the usual lighter-than-water glass sphere. This outer vessel is closed at the top by a rubber bung, through a hole in which a sample of the dilute solution may be removed with a pipette. The results, obtained at 60°, are given in Table V.

TABLE V. DIFFUSION FROM ELASTIC SOLUTION AT 60° C

K cetane sulphonate conc.	KCl conc.	$\bar{D} \times 10^7$	
		G 3	G 4
0.01	0.04	9.4	0.11
		8.6	0.11
		—	0.07

It will be seen that there is now an enormous discrepancy between the values given by the two diaphragms, in the same sense as that found previously but altogether different in magnitude. It is almost certain that the stirring on the "concentrated" side of the diaphragm was inadequate, but there is no reason to suppose that it was any less adequate for the G4 diaphragm. The hundred-fold reduction of diffusion velocity through the finer diaphragm must be explained almost entirely by the internal structure which can presumably build up more completely in this case. This very striking result may seem to be in contradiction to the conclusion of Laing-McBain (1933), who found no appreciable difference between the rate of diffusion from a mobile sodium oleate solution and a clear gel of the same concentration. It must be remembered that diffusion in this case was allowed to occur into water so that the transfer of paraffin-chain micelles was mainly due to the diffusion potential. It might not therefore be affected by the building up of larger structures since we are concerned with their electrical rather than their diffusion mobility.

Diffusion of micelles containing other molecules

Aqueous solutions of paraffin-chain salts are able to dissolve considerable amounts of many organic substances which are almost insoluble in water. The organic substance dissolves in the micelles (Smith 1932; Hartley 1937, p. 153; Lawrence 1937). Some diffusion measurements in such systems were made with a view to determining how much the micelle is swollen by the additional molecules.

The measurements were made with cetyl pyridinium chloride and with choline cetane sulphonate, sodium and choline chlorides being used as swamping electrolytes. The organic liquids used, benzene and *n*-amyl alcohol, were added slowly to the aqueous solutions from a microburette until saturation was reached. The equilibrium was reached very slowly with benzene, but almost immediately with the alcohol. The saturation concentrations were very sensitive to temperature, and the room was therefore maintained at 25° C during the additions. Although, when half saturated, the solutions were very viscous, when fully saturated they became normally mobile (Angelescu and Popescu 1930). They showed a pronounced Tyndall cone, though similar solutions without the swamping electrolyte are clear.

The rotational cells were used at 25° C. Although cetyl pyridinium chloride is precipitated by the sodium chloride at this temperature in the absence of the organic liquid, the presence of the latter lowers the solution temperature to well below its value in the absence of the sodium chloride.

In order to guard against loss of the benzene by evaporation, a layer of the free liquid was carefully run on top of the "dilute" solution in the cell. This precaution was omitted in the experiments with alcohol, since it was less necessary with the less volatile liquid and would have introduced errors owing to the high solubility of the paraffin-chain salts in the free alcohol.

TABLE VI

Paraffin-chain salt and conc.	Swamping salt and conc.	Molecular ratio in micelle: organic liquid to paraffin chains	$\overline{D} \times 10^7$	
			G 3	G 4
Cetyl pyridinium chloride	Sodium chloride	Benzene		
0.05	0.5	0	8.5	8.5
0.05	0.5	4	13.3	5.2
0.05	1.0	4	14.5	—
		Amyl alcohol		
0.05	1.0	3.8	3.4	0.8
Choline cetane sulphonate	Choline chloride	Benzene		
0.05	1.0	0	6.0	6.5
0.05	1.0	4.4	5.1	3.6

The results are shown in Table VI. For comparison, the diffusion coefficients measured in the presence of the swamping electrolyte alone are also given. It will be seen that not only is there a large discrepancy between the values obtained with the diaphragms of different porosity but that in the case of the benzene-cetyl pyridinium chloride solutions, the diffusion coefficient through the G 3 diaphragm has *increased* by some 50 %. The other solutions give the expected lower diffusion coefficient, although the discrepancy between the two diaphragms prevents any conclusion being drawn as to the micelle size.

The presence of very small specks of benzene or amyl alcohol occasionally observed on examining the diaphragms after the diffusion measurement suggests a possible explanation of these anomalous results. Dilute solutions of cetyl pyridinium chloride leave glass vessels in a peculiarly greasy condition, due to the adsorption of the cetyl pyridinium cations on the negatively charged glass walls, leaving the paraffin part exposed to the solution. This paraffin layer is, however, wetted by the solutions themselves unless they are very dilute. In the diffusion measurements there will thus

be a layer of paraffin lining the walls of the capillaries in the porous diaphragm. As the thickness of this layer (of the order of the length of the paraffin chain) is small compared with the diameter of the capillaries, there will normally be no effect upon the rate of diffusion, and this is reflected in the results with cetyl pyridinium chloride discussed above. In the case of the benzene-cetyl pyridinium chloride solutions, however, there is the possibility of a much thicker layer of *benzene* becoming attached to the walls of the capillaries through the medium of the adsorbed cetyl pyridinium ions. In the smaller capillaries the benzene layer might be sufficiently thick to block them completely, thus reducing the effective diffusing area of the diaphragm. The benzene layer, however, would also provide a medium through which the cetyl pyridinium chloride could diffuse with an increased velocity. The complete immersion of the cetyl pyridinium cations in the benzene is improbable, but by being adsorbed at the benzene-solution interface, the ions could move through the capillary with their polar groups in the aqueous solution and the paraffin chains in the benzene. The cations, having passed through the capillary, could then reform into micelles. Thus the expected reduction in the rate of diffusion, due to the increase in micelle size upon solution of the benzene, would be further reduced by the blocking of some of the smaller capillaries in the diaphragm, while the partial blocking of some of the larger capillaries would result in an accelerated "surface diffusion". The observed rate of diffusion would be governed by these three factors, but would probably be lower in the diaphragm with the smaller pore size.

It might be expected that the amyl alcohol, having more affinity for water than has benzene, would be less adsorbed on the walls of the capillaries and that the surface diffusion would therefore be reduced. This is partly borne out by the results, for, although there is still a discrepancy between the G3 and G4 diaphragms, there is no acceleration of diffusion above its normal value.

That the presence of the layer of organic liquid on the capillary walls is dependent on the initial adsorption of paraffin-chain ions is suggested by the better agreement between the G3 and G4 diaphragms when choline cetane sulphonate was used as the solvent for the benzene or alcohol. The accelerated surface diffusion does not then predominate in either diaphragm. The adsorption of the paraffin-chain anions on the negative glass walls is, of course, less than that of the paraffin-chain cations.

Although it is possible that, with lower concentrations of organic solute, a better agreement, and therefore more significant results, might have been obtained, the results suggest that recourse must be had to the free liquid

method of diffusion measurement for these systems, as for the sulphonates and sulphates. For such slow diffusing substances, however, very great precautions against convection would be necessary.

It is a pleasure to acknowledge our indebtedness to Professor F. G. Donnan for his interest and encouragement, and to the Research Council of Imperial Chemical Industries for a grant which made the investigation possible. To Dr Conmar Robinson we are indebted for helpful discussion of some of the problems that were encountered during the course of the work.

SUMMARY

1. The porous diaphragm method has been applied to the measurement of the diffusion of paraffin-chain micelles in aqueous solution with a swamping excess of simple electrolyte.

2. For cetyl pyridinium salts the method is satisfactory and a radius for the micelle of about 26 Å is found. The radius is independent, within experimental error, of the concentration of paraffin chain or swamping electrolyte, and very little dependent on temperature. These results are in excellent agreement with the "spherical liquid micelle" theory. There is small, but definite, influence of the nature of the gegenion on the size of the micelle.

3. For cetane sulphonates and cetyl sulphates the method fails to give reliable results, probably owing to the building up of a structure within the pores of the diaphragm.

4. When the solutions are saturated with amyl alcohol or benzene the method again fails to give reliable results. A mechanism is suggested for the anomalies observed.

REFERENCES

- Adam, N. K. 1936 *Ann. Rep. Chem. Soc.* for 1936, **33**, 103-17.
Adam, N. K. and Shute, H. L. 1938 *Trans. Faraday Soc.* **34**, 758-766.
Angelescu, E. and Popescu, M. 1930 *Kolloidzshr.* **51**, 336-48.
Bruins, H. R. 1931 *Kolloidzshr.* **54**, 265-78; **57**, 152-166.
— 1932 *Kolloidzshr.* **59**, 263-6.
Bury, C. R. and Jones, E. R. 1927 *Phil. Mag.* **4**, 841-8.
Bury, C. R. and Grindley, J. 1929 *J. Chem. Soc.* pp. 679-84.
Bury, C. R. and Davies, D. G. 1930 *J. Chem. Soc.* pp. 2263-7.
Freundlich, H. and Kruger, D. 1935 *Trans. Faraday Soc.* **31**, 906-13.
Hartley, G. S. 1934 *Trans. Faraday Soc.* **30**, 444-50.

- Hartley, G. S. 1935 *Trans. Faraday Soc.* **31**, 257-8.
— 1936a "Aqueous Solutions of Paraffin-Chain Salts." Paris: Hermann.
— 1936b *J. Amer. Chem. Soc.* **58**, 2347-54.
— 1937 Article in "Wetting and Detergency", pp. 153-62. London: Harvey.
— 1938a *Trans. Faraday Soc.* **34**, 1283-8.
— 1938b Paper on solubility of azobenzene communicated to *J. Chem. Soc.*
Hartley, G. S., Collie, B. and Samis, C. S. 1936 *Trans. Faraday Soc.* **31**, 795-815.
Hartley, G. S. and Robinson, Conmar 1931 *Proc. Roy. Soc. A*, **134**, 20-35.
Howell, O. R. and Robinson, H. G. B. 1936 *Proc. Roy. Soc. A*, **155**, 386-406.
Howell, O. R. and Warne, H. 1937 *Proc. Roy. Soc. A*, **160**, 440-54.
Jander, G. and Weitendorf, K. F. 1934 *Angew. Chem.* **47**, 197-200.
Kohlrausch, F. and Holborn, L. 1916 "Das Leitvermögen der Elektrolyte." Leipzig: Teubner.
Lawrence, A. S. C. 1937 *Trans. Faraday Soc.* **33**, 815-20.
Laing-McBain, M. E. 1933 *J. Amer. Chem. Soc.* **55**, 545-51.
McBain, J. W. and Liu, T. H. 1931 *J. Amer. Chem. Soc.* **53**, 59-74.
McBain, J. W. and Woo, F. M. 1938 *J. Amer. Chem. Soc.* **60**, 223-7.
Meyer, K. H. and van der Wyk, A. 1937 *Helv. chim. Acta*, **20**, 1321-30.
Murray, R. C. and Hartley, G. S. 1935 *Trans. Faraday Soc.* **31**, 183-9.
Powney, J. and Addison, C. C. 1937 *Trans. Faraday Soc.* **33**, 1243-53.
Samis, C. S. and Hartley, G. S. 1938 *Trans. Faraday Soc.* **34**, 1288-1305.
Smith, E. Lester 1932 *J. Phys. Chem.* **36**, 1401-18, 1672-84.
Thiessen, P. A. and Spychalski, R. 1931 *Z. phys. Chem. A*, **156**, 435-56.
Valko, E. 1936 *Trans. Faraday Soc.* **31**, 230-44.
-

The thermal decomposition of nitrous oxide

By R. M. LEWIS AND C. N. HINSHELWOOD, F.R.S.

(Received 15 July 1938)

When the thermal decomposition nitrous oxide was first investigated reactions were usually thought of as belonging to simple integral orders. Over the range 100–700 mm. the nitrous oxide reaction proved to be more nearly of the second than of any other order, showing that the activation process was collisional (Hinshelwood and Burk 1924), and not one depending upon the absorption of radiation by isolated molecules—a possibility at one time considered. For a first-order reaction the plot of the reciprocal of the half-reaction time against pressure is a line parallel to the pressure axis, while for a second-order reaction it is a line inclined to the axis and passing through the origin. In a series of investigations (Volmer and Kummerow 1930; Volmer and Nagasako 1930; Musgrave and Hinshelwood 1932; Hunter 1934) it has come to light that the form of this curve for nitrous oxide is really rather complex, and may be divided into the following parts: (*a*) an initial steep increase, starting from the origin, which between 50 and 100 mm. becomes shallower and passes into (*b*) an almost linear curve continuing up to several atmospheres when it gradually bends again passing into (*c*) an almost straight line of still smaller slope which continues up to 20–30 atm. when it bends round and becomes nearly parallel to the pressure axis, as for a reaction of the first order. Various interpretations have been suggested: (1) The curve results from the superposition of three “quasi-unimolecular” reactions, each of the second order at low pressures and of the first order at higher pressures (typical of reactions in which activation is by collision and is followed by transformation of isolated molecules). The three components represent different activation modes with different transformation probabilities of the activated molecules (Hinshelwood, Fletcher, Verhoek and Winkler 1934). (2) The curve is not really composite in form, but represents a single quasi-unimolecular reaction, the transformation probability of the molecules being a continuous function of the excess energy they contain. This view was supported by Volmer who, however, was not aware of the existence of region (*c*) of the curve, and believed that the portion (*b*) became parallel to the axis above about 10 atm. He did not recognize the distinctness of portion (*a*) but plotted the reciprocal of the velocity constant against the reciprocal of the pressure so that points corresponding to low pressures were spread out into an indefinite sweep which masked

the normal composite appearance of the curve. (3) The third interpretation has not been specifically formulated for the case of nitrous oxide, but is implied by Letort's treatment (Letort 1937) of the analogous example of acetaldehyde: it amounts to the view that the changes of slope shown by the curve are not so much due to changes from one integral reaction order to another as to the existence of fractional orders, such as the order $3/2$ which arises in certain circumstances when the mechanism depends upon the intervention of atoms or radicles. Limited stretches of the nitrous oxide curve could be represented approximately by the equation of a reaction of the $3/2$ order: and, although this would not apply over the whole range, nevertheless, if we assume that the curve is complex, we should not neglect the possibility that one of the components is of this type. The present paper contains new experimental data bearing upon these interpretations, and upon other problems presented by this interesting reaction.

EXPERIMENTAL DETAILS

The reaction is $2\text{N}_2\text{O} = 2\text{N}_2 + \text{O}_2$, with a very small proportion of $2\text{N}_2\text{O} = 2\text{NO} + \text{N}_2$, each of which corresponds to a 50 % increase of pressure.

The reaction was followed by measuring the change of pressure with time, in a 300 c.c. silica bulb, heated in an electric furnace, and provided with a manometer for direct reading, and with a Bourdon gauge which was sensitive enough to give reliable readings of the initial rate of reaction. The Bourdon gauges used were calibrated directly against the manometer, the relationship between the deflexion reading in the microscope eyepiece and the pressure being shown to be linear. The temperature was measured by a Pt/Pt-Rh thermocouple and a pyrometer, the pointer of which was read by a microscope so that variations of 0.1° could be detected. Special care was taken to keep the temperature constant, and to reproduce it accurately from day to day. Since the accuracy with which the temperature could be controlled during a few hours was greater than that with which it could be reproduced, the eyepiece scale was standardized each day with the aid of a rate measurement for pure nitrous oxide at a standard pressure. This corrected for the slight instability of the pyrometer zero, and ensured that the relative temperatures were accurately controlled: they should be within 0.5° . The absolute temperature was calculated after the pyrometer had been calibrated in terms of the melting-point of pure antimony, and should be within about 2° .

Nitrous oxide was obtained from a cylinder and the only impurities were oxygen and nitrogen: it was condensed in a trap surrounded by liquid air,

and all the permanent gases were pumped away before the nitrous oxide was allowed to evaporate into its container. Carbon dioxide was made by heating sodium bicarbonate; nitrogen and argon were obtained from cylinders and dried.

When working with mixtures of nitrous oxide with other gases, the separate gases were introduced into a pre-mixer before admission to the reaction bulb.

The dead space was estimated as accurately as possible and the appropriate corrections applied to the results for runs followed manometrically and runs followed by the Bourdon gauge, the dead space being increased in the latter case. When the temperature difference had been allowed for, these corrections were equivalent to 7 and 9 % of the bulb volume, and they were applied in a consistent way in all the experiments, so that if the half-reaction point was estimated in error as $(50-d)\%$, d should be the same throughout.

Several end-points were determined and when this correction had been applied the mean increase in pressure was 49 % of the initial pressure.

Half-reaction time and initial rate of reaction

The interpretation referred to above under (1) has been criticized on the ground that it is based upon plots of the reciprocal half-time instead of an initial rate. If the differential equation for the reaction rate is a sum of several terms, it is evident that these terms can not be integrated independently, and that the reciprocal half-time will not be the sum of three reciprocal half-times characteristic of the separate components. It should, however, be obvious that if the integral curve is composite in form, this fact can hardly be due to anything but the composite nature of the differential expression from which it is derived. Nevertheless, it seemed desirable to see how far the curve of reciprocal half-time against pressure differed from that of the initial rate. The initial rates were therefore determined with care for comparison with the half-times. From Table I and the figs. 1 and 2 it appears that the general course of the $1/(\text{half-time})$ curve is closely similar to that of the curve for $1/p \cdot dp/dt$, p being the initial pressure, and dp/dt the initial rate. It is clear, therefore, that a distortion of the curve based upon half-times, in such a way as to simulate a composite nature which is not really inherent in the differential equation for the reaction rate, is out of the question. In figs. 1 and 2 the actual values of $1/p \cdot dp/dt$ are plotted, while the values of $1/t_{\frac{1}{2}}$ are multiplied by a constant to make them equal to the former at 300 mm.

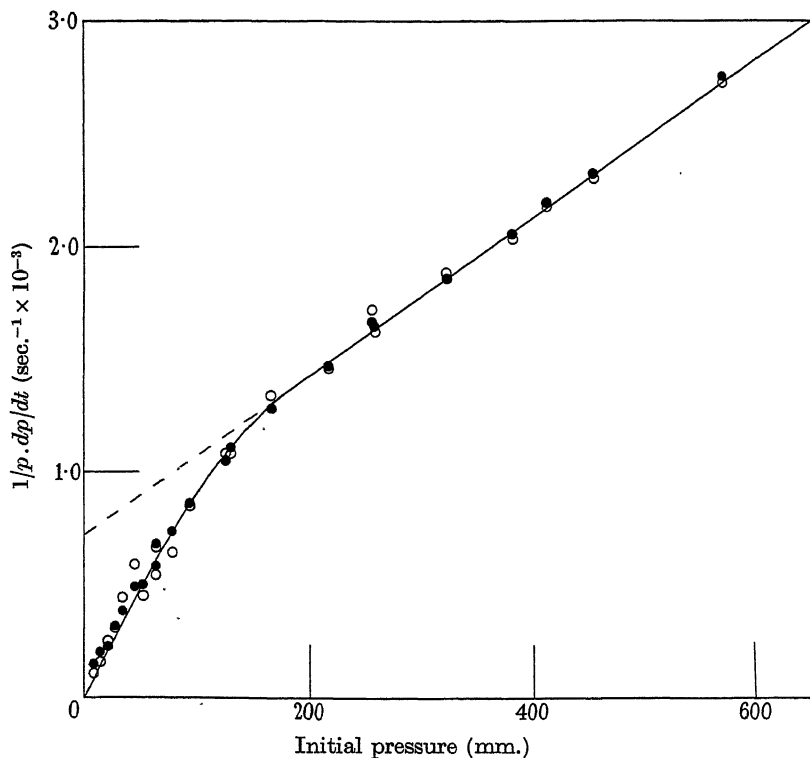


FIG. 1. Influence of pressure on reaction rate at 747° C, showing similarity of curves derived from half-times and initial rates respectively. Circles = $1/p \cdot dp/dt$; black dots = $1/t_4 \times \text{constant}$.

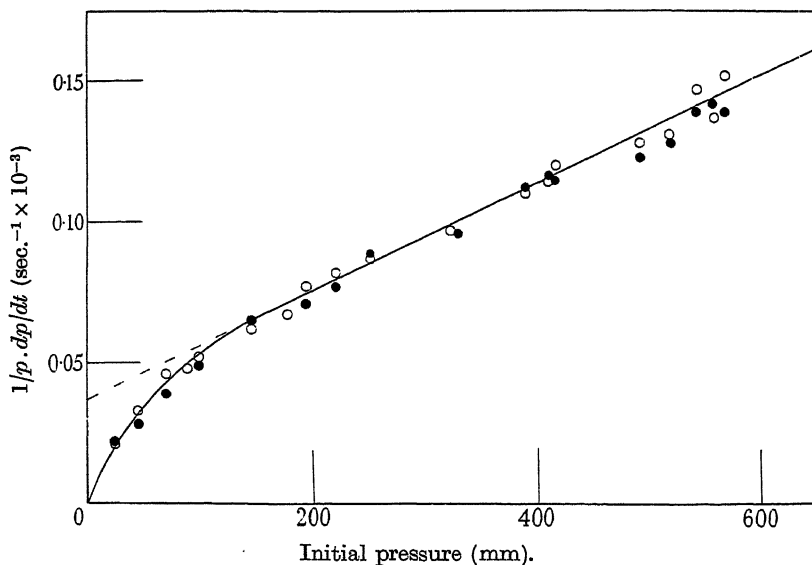


FIG. 2. Influence of pressure on reaction rate at 652° C. Circles = $1/p \cdot dp/dt$; black dots = $1/t_4 \times \text{constant}$.

TABLE I

N ₂ O—747° C			N ₂ O—652° C		
<i>p</i> mm.	1/ <i>t</i> _i sec. ⁻¹ × 10 ⁻⁵	1/ <i>p</i> · <i>dp/dt</i> sec. ⁻¹ × 10 ⁻³	<i>p</i> mm.	1/ <i>t</i> _i sec. ⁻¹ × 10 ⁻⁵	1/ <i>p</i> · <i>dp/dt</i> sec. ⁻¹ × 10 ⁻³
8.9	47	0.11	24.4	6.9	0.022
15.0	65	0.16	45.3	8.7	0.033
22.5	72	0.25	70.2	12.1	0.046
27.8	98	0.31	88.2	—	0.048
35.7	120	0.44	99	15.3	0.052
46.0	156	0.59	147.5	20.3	0.062
53.3	159	0.45	179	—	0.067
64.5	216	0.64	195	22.4	0.077
65.0	186	0.54	222	24.1	0.082
79.3	236	0.64	252	28.0	0.087
95	276	0.85	324	—	0.097
127	339	1.05	330	30.1	—
127	336	1.08	390	35.2	0.110
131.5	355	1.08	410.5	36.5	0.115
166	407	1.34	415.5	36.0	0.120
217	465	1.46	492	38.6	0.128
256	532	1.72	519	40.0	0.131
258	524	1.63	543	43.7	0.147
322	595	1.88	557	44.6	0.137
382	658	2.03	568	43.5	0.152
412	699	2.18			
452	740	2.30			
569	877	2.72			

*Variation of activation energy with pressure. Determinations
from half-times and from initial rates*

The initial rate-initial pressure curve as well as the curve for 1/(half-time) initial pressure was determined for two temperatures nearly 100° apart. From the ratios at different pressures the activation energies were calculated by the Arrhenius equation. The results confirm the previously reported increase in activation energy with increasing pressure. The meaning of this result is that at the lower pressures molecules with smaller energies are contributing a greater proportion to the total of the observed change. If the probability of transformation is smaller for the lower energies, then it is only at the lower pressures that the molecules with the lower activation energies survive deactivation by collision long enough to contribute to the observed reaction, so that the result is easily understandable. It does not, however, tell us anything about the problem as to whether there is a continuous variation of transition probability with energy content or whether there are distinct sets of activated molecules which differ in kind. If the reality of the composite rate-pressure curve is accepted, then we have

the latter possibility. The actual values of the activation energy agree fairly closely with those given by Hinshelwood and Burk (1924), Musgrave and Hinshelwood (1932) and by Hunter (1934): they are considerably higher than those given by Volmer (Volmer and Kummerow 1930; Volmer and Nagasako 1930). Upon the low values reported by the latter author various theoretical considerations about the non-adiabatic nature of the nitrous oxide decomposition have been based. In the light of the higher values these considerations lose their significance (Wigner 1938).

The activation energies are tabulated in Table II. Those derived from the half-times are slightly greater at any given pressure. In each case the change with pressure follows a similar curve.

TABLE II. ACTIVATION ENERGIES GIVEN BY $E = \frac{1.98 \times 2.303}{1/T_2 - 1/T_1} \times \log_{10} r$ cal.,

WHERE r IS THE RATIO OF $1/t_{\frac{1}{2}}$ OR OF $1/p \cdot dp/dt$ AT THE TWO TEMPERATURES T_1 AND T_2

$T_1 = 747^\circ \text{C}$; $T_2 = 652^\circ \text{C}$

Pressure mm.	r from $1/t_{\frac{1}{2}}$	E	r from $1/p \cdot dp/dt$	E
25	16.0	54,500	11.3	47,700
50	17.4	56,200	13.4	51,000
75	17.8	56,600	15.1	53,400
100	18.2	57,100	16.6	55,300
150	18.7	57,600	18.2	57,100
200	18.8	57,700	18.8	57,700
300	19.2	58,000	18.8	57,700
400	19.4	58,300	18.8	57,700
500	19.7	58,600	18.8	57,700
600	19.8	58,700	18.8	57,700

Influence of the reaction products

It is known that small quantities of nitric oxide are produced in the reaction and that these exert a catalytic effect: it is also known that inert gases, such as nitrogen and oxygen, exert some accelerating effect. To show the magnitude of this influence Table III records a complete pressure-time curve for a reaction in which the initial pressure was fairly great. From this curve the velocity over small intervals was read off at different points of the curve, for which also the mean nitrous oxide pressure and the mean pressure of products could be calculated. The rather abrupt finish of the reaction is to be noted: the last traces of nitrous oxide in presence of a large accumulation of products decomposing with greatly increased speed. Fig. 3 shows the influence of the products in preventing the fall in reaction rate at low partial pressures of nitrous oxide. Under the conditions of this

TABLE III. INITIAL PRESSURE OF NITROUS OXIDE = 402 MM.

Time sec.	Change in pressure	Time sec.	Change in pressure
—	—	161	88
0	0	175	93
5	3	193	98
7.5	5	209	103
11	8	226	108
17	13	253	113
24	18	278	118
32	23	304	123
37	28	333	128
48	33	368	133
55	38	404	138
63	43	447	143
73	48	504	148
81	53	570	153
90	58	655	158
100	63	780	163
111	68	990	168
123	73	1560	172
135	78	1860	173
148	83	54 min. (inf.)	173.3

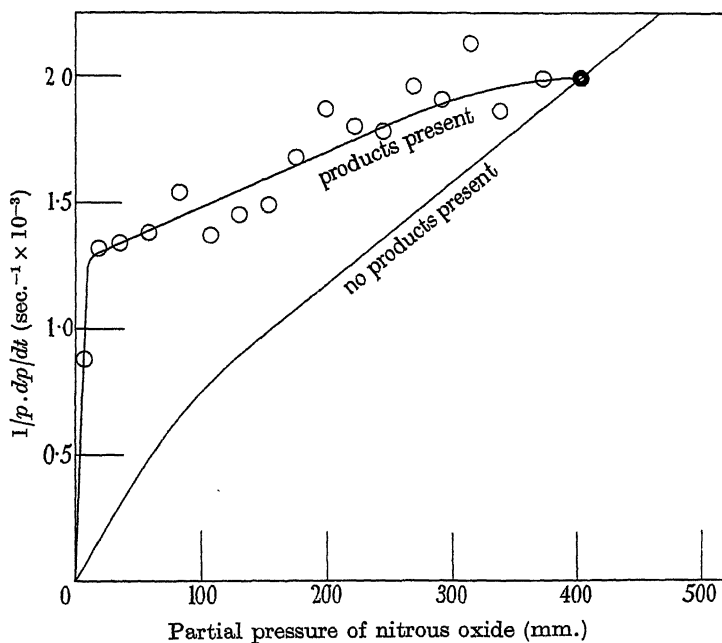


FIG. 3. Influence of reaction products on the curve of initial rate against nitrous oxide pressure. The rates are found by drawing tangents to a curve plotted from the results given in Table III. The lower curve is reproduced from fig. 1.

experiment the excess of products becomes very large for the smaller partial pressures of nitrous oxide: e.g. when the latter has fallen to 10 mm. there will be about sixty times this pressure of products.

Influence of inert gases on the reaction rate

Volmer and Bogdan (1933) determined the influence of various inert gases on the reaction rate. They believed that the reaction consisted of one single quasi-unimolecular reaction, whose unimolecular constant attained its limiting rate in the neighbourhood of 10 atm., and that the influence of the inert gases was to help the constant for lower pressures to rise towards this limiting rate. They found this limiting rate by an extrapolation which the experiments of Hunter have since shown, to be unjustified, and they plotted their results for the inert gas influence in such a way as to force an extrapolation to this same limit. Since this limit has no real existence the curves they give must be regarded as largely arbitrary.

TABLE IV

p = partial pressure of N_2O at 652°C

$\text{N}_2\text{O} + 2\text{CO}_2$		$\text{N}_2\text{O} + 2\text{N}_2$		$\text{N}_2\text{O} + 2\text{A}$		$\text{N}_2\text{O} + 4\text{A}$	
p	$1/p \cdot dp/dt$	p	$1/p \cdot dp/dt$	p	$1/p \cdot dp/dt$	p	$1/p \cdot dp/dt$
25	0.052	20	0.028	22	0.029	25	0.035
45	0.072	29	0.041	30	0.031	40	0.049
63	0.080	49	0.054	34	0.043	50	0.055
82	0.092	59	0.059	39	0.041	60	0.057
90	0.100	73	0.065	50	0.048	74	0.066
101	0.113	80	0.068	61	0.050	78	0.062
126	0.117	100	0.076	75	0.059	90	0.073
150	0.131	125	0.082	92	0.065	102	0.079
176	0.135	148	0.096	100	0.070	110	0.084
200	0.144	150	0.096	125	0.074	115	0.083
307	0.180	175	0.100	152	0.082	145	0.092
330	0.182	200	0.114	175	0.087	182	0.096
361	0.216	227	0.114	200	0.093	200	0.104
423	0.218	245	0.125	221	0.093	216	0.097
424	0.220	299	0.137	264	0.108	217	0.101
458	0.231	310	0.129	298	0.107	230	0.111
492	0.272	349	0.139	342	0.120	286	0.120
		384	0.143	352	0.117	327	0.130
		396	0.156	401	0.131	339	0.146
		426	0.171	432	0.139	359	0.136
		471	0.172	486	0.143		
		517	0.190	490	0.147		
		566	0.191	564	0.165		

Experiments have been made on the initial rates of reaction for the following mixtures: $\text{N}_2\text{O} + 2\text{CO}_2$, $\text{N}_2\text{O} + 2\text{N}_2$, $\text{N}_2\text{O} + 2\text{A}$, $\text{N}_2\text{O} + 4\text{A}$, the total

initial pressure being varied over a range of about two atmospheres. The results are given in Table IV and figs. 4 and 5. With carbon dioxide the increase in rate caused by the inert gas is almost uniform over the pressure range employed, the acceleration tending if anything to diminish at the higher total pressures. With nitrogen and with argon the influence is greatest

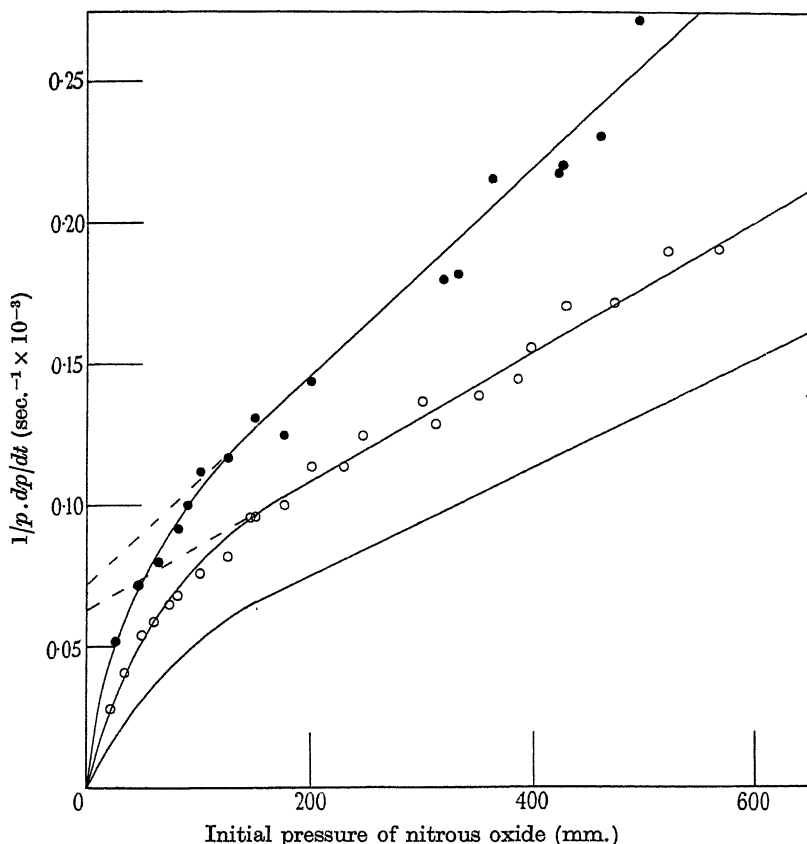


FIG. 4. Influence of carbon dioxide and of nitrogen at 652°C. Black dots, $\text{N}_2\text{O} + 2\text{CO}_2$; circles, $\text{N}_2\text{O} + 2\text{N}_2$; lower line, N_2O alone.

when the partial pressure of the nitrous oxide is lowest as is shown in figs. 4 and 5 which reveal the fact that the influence of the inert gas is exerted principally on the low-pressure component of the total reaction. At the higher pressure the curves representing the rate in presence of argon become almost parallel to the normal curve, which is not at all what would be expected if the inert gas were contributing to the activation of the molecules in one single quasi-unimolecular reaction (Volmer interpretation). With

the assumption of more than one unimolecular reaction it is easy to explain the variable effect of the foreign gases, since transfer of activation energy is a specific phenomenon and may well be different for different types of activated state. This explanation, however, also meets with difficulties. According to the theory of the composite mechanism, the reaction over the range of pressure used in these experiments consists of one reaction which completes its change from the second order to the first, and a second part which is of the second order up to much higher pressures (the second part

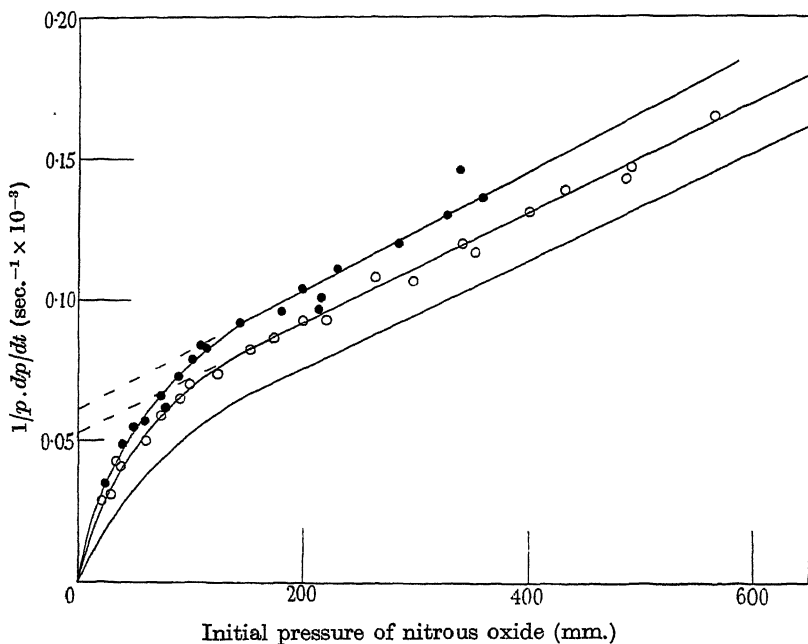


FIG. 5. Influence of argon at 652° C. Black dots $\text{N}_2\text{O} + 4\text{A}$; circles $\text{N}_2\text{O} + 2\text{A}$; lower line N_2O alone.

being itself composite but representable by one constant in this range). The rate of reaction will be given by

$$-\frac{1}{p} \cdot \frac{dp}{dt} = \frac{k_1 p}{1 + bp} + k_2 p$$

for nitrous oxide alone, and by

$$-\frac{1}{p} \cdot \frac{dp}{dt} = \frac{k_1 p + k'_1 P}{1 + bp + b'P} + k_2 p + k'_2 P$$

in the presence of the foreign gas the pressure of which is P . In the experiments P always bore a definite ratio to p , the pressure of the nitrous oxide.

Further, if the mechanism of the foreign gas effect is one of simple activation and deactivation, the ratio of k'_1 to k_1 should be the same as that of b' to b (principle of microscopic reversibility). The equation then assumes the form

$$-\frac{1}{p} \cdot \frac{dp}{dt} = \frac{\gamma k_1 p}{1 + \gamma b p} + \Gamma k_2 p.$$

This equation reproduces the results excellently for any given curve, a fact to which no great weight can be attached since there are three constants. The ratio of k_1 to b should be constant. This requirement is, however, not verified, as the following numbers show.

Mixture	γk_1	γb	Γk_2	k_1/b
N ₂ O	1.75×10^{-3}	4.73×10^{-2}	1.92×10^{-4}	0.037
N ₂ O + 2CO ₂	5.47×10^{-3}	7.60×10^{-2}	3.68×10^{-4}	0.072
N ₂ O + 2A	2.69×10^{-3}	5.04×10^{-2}	1.94×10^{-4}	0.053
N ₂ O + 4A	4.15×10^{-3}	6.80×10^{-2}	2.10×10^{-4}	0.061
N ₂ O + 2N ₂	3.14×10^{-3}	4.98×10^{-2}	2.26×10^{-4}	0.063

The figures in the last column show that, if we regard the lower part of the curve as consisting largely of a unimolecular reaction which attains its limiting rate in the region where the whole curve bends rapidly round, then this limiting rate itself is raised by the presence of the foreign gas. Fig. 6 shows the components into which the total curves can be analysed in accordance with the above equation, and reveals the increase by the inert gas of the limiting "low-pressure" rate. This result shows that the method of representation is incorrect or incomplete, even though it expresses the velocities correctly within the limit of experimental error. We therefore reconsider the interpretations (1), (2) and (3) referred to above. (2) and (3) are difficultly reconcilable with the results of Hunter (1934) for the whole range of pressure of 40 atm. which has been investigated, but they might be applicable to the range here considered, leaving alternative activation modes to account for the high-pressure results. Even for the low-pressure range, however, they have their difficulties. As regards (2), it would account only with difficulty for the rapid change of direction of the curve in the region of 100 mm. (the masking of this in Volmer's method of plotting has been referred to). It would not account for the fact that with nitrogen and argon the constant Γk_2 changes so little with addition of the foreign gas compared with γk_1 . With regard to (3) it must be remembered that in the unimolecular decomposition oxygen atoms are produced: if these caused catalytic decomposition of more nitrous oxide, and if they recombined at a rate proportional to the square of their concentration, then the expression for the reaction rate, although it would be rather complicated in actual form,

could easily approximate to one involving $[\text{N}_2\text{O}]^\ddagger$, which would account for the change in direction of the curve in the neighbourhood of 100 mm. (A superposition of a curve of this form on the curve of a quasi-unimolecular reaction for higher pressures could by suitable manipulation of constants probably be made to reproduce the results with moderate accuracy. But this view entails further difficulties. In the first place the action of oxygen atoms on nitrous oxide has already been invoked to account for the production of nitric oxide $\text{N}_2\text{O} + \text{O} = 2\text{NO}$, and the proportion of this found is small (Musgrave and Hinshelwood 1932). And secondly, it is difficult to see why the inert gas should not help the recombination of the oxygen atoms and so retard rather than accelerate the reaction.)

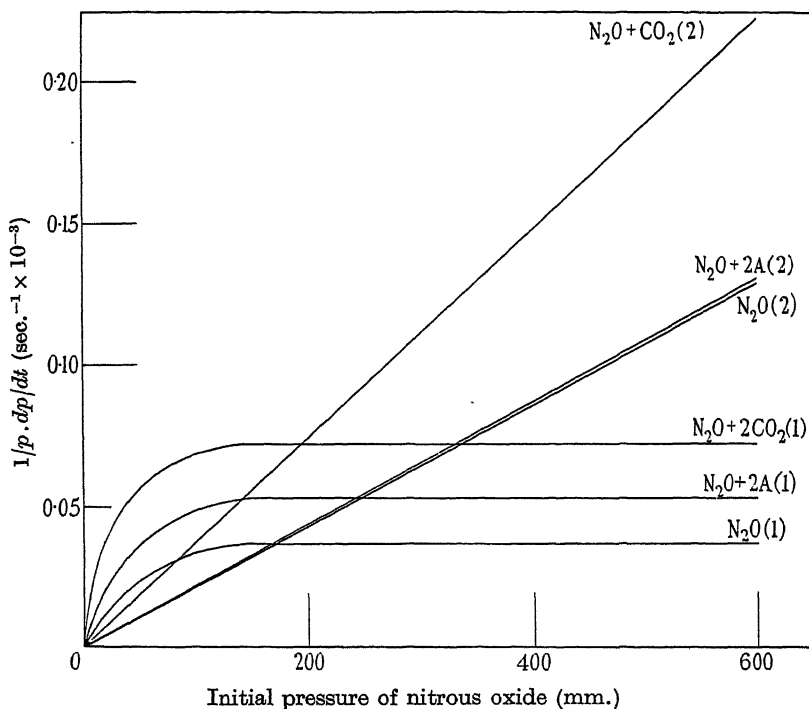


FIG. 6. Formal analysis of curves into possible components. (1) In each case is the hypothetical low-pressure quasi-unimolecular component, and (2) the component which remains of the second order over this range of pressure.

It remains to consider the following hypothesis: carbon dioxide, in virtue of its close similarity in structure to nitrous oxide, is capable of activating nitrous oxide molecules in the same ways as they activate one another in collisions, as is shown by the marked increases in γk_1 and Γk_2 . With nitrogen

and argon, however, Γk_2 is not seriously affected, while γk_1 increases, though to a less extent than with carbon dioxide. As k_1/b is not constant, it must be supposed that all three gases are capable of producing by their collisions a form of motion in the activated nitrous oxide molecule which is not produced in normal nitrous oxide collisions, and which is associated with a small transformation probability, as well as the normal activated form. The effect of this will be to add another, more or less independent, quasi-unimolecular element to the composite mechanism which will be prominent at the lower pressures and will have the required result of raising the low pressure part of the curve in the required manner. The equation would be

$$-\frac{1}{p} \cdot \frac{dp}{dt} = \frac{\gamma k_1 p}{1 + b p} + \Gamma k_2 p + \frac{k_1'' p}{1 + b'' p},$$

k_1/b no longer has to be constant, but can rise to

$$\left(\frac{k_1}{b} + \frac{k_1''}{b''} \right)$$

as found.

It is quite possible to analyse the composite curve found in presence of the foreign gas into the components which this interpretation demands: but the value of such a procedure is doubtful, since fresh constants have to be introduced, and, moreover, it is difficult to make allowance for the effect of the oxygen atoms at low pressures, and the production of the small amount of nitric oxide shows that some effect must be present. The further increase of k_1/b on increasing the proportion of argon from 2 to 4 also shows that such an analysis can not give by itself a complete explanation. Before, therefore, the exact mechanism of the action of the foreign gases at low pressures can be settled conclusively fresh evidence of a qualitatively new kind is probably needed.

The result of this part of the present investigation may therefore best be stated in the following form. At the higher pressures (100–600 mm.) of nitrous oxide the foreign gases, argon and nitrogen, do not contribute much to the activation of the molecules principally concerned in the reaction. Carbon dioxide has, however, an effect comparable with that of nitrous oxide itself. At the lower pressures all the gases have an influence in facilitating reaction, whether because they are specifically capable of activating nitrous oxide molecules to states of relatively low activation energy with small transformation probabilities, or whether partly or wholly for another reason.

The contrast between the case of argon and nitrogen which are mainly effective in the low-pressure region and carbon dioxide which is nearly as

effective in the high-pressure region also, is further evidence that the reaction mechanism is kinetically composite.

SUMMARY

Further experimental data relating to the thermal decomposition of nitrous oxide at 747° and 652° C are recorded.

The variation of reaction rate with pressure is expressed both in terms of the reciprocal half-reaction time and in terms of the initial rates. The two curves obtained are similar enough to show that previous conclusions based upon the half-time curve are confirmed by the initial rate curve.

The mean value of the activation energy falls at lower pressures, and the absolute magnitude of this energy agrees with most previous determinations, but not with those upon which has been based a theory that the reaction is abnormally slow for certain quantum-mechanical reasons.

The influence of the products of reaction and of additions of nitrogen, argon and carbon dioxide have been studied. There is a contrast between the action of argon and that of carbon dioxide in that the former has a relatively greater effect at lower than at higher partial pressures of nitrous oxide. This fact is consistent with a greater complexity of the reaction mechanism than corresponds to the assumption of a single quasi-unimolecular reaction. The nature of this complexity is discussed.

REFERENCES

- Hinshelwood and Burk 1924 *Proc. Roy. Soc. A*, **106**, 284.
Hinshelwood, Fletcher, Verhoek and Winkler 1934 *Proc. Roy. Soc. A*, **146**, 327.
Hunter 1934 *Proc. Roy. Soc. A*, **144**, 386.
Letort 1937 *J. Chim. phys.* **34**, 206, 265, 355.
Musgrave and Hinshelwood 1932 *Proc. Roy. Soc. A*, **135**, 23.
Volmer and Bogdan 1933 *Z. phys. Chem. B*, **21**, 257.
Volmer and Froelich 1932 *Z. phys. Chem. B*, **19**, 85.
Volmer and Kummerow 1930 *Z. phys. Chem. B*, **9**, 141.
Volmer and Nagasako 1930 *Z. phys. Chem. B*, **10**, 414.
Wigner 1938 *Trans. Faraday Soc.* **34**, 38, 70.
-

Progressive lightning. VI

BY B. F. J. SCHONLAND, F.R.S.,* D. J. MALAN, *Docteur de l'Université de Paris*† AND H. COLLENS, M.I.E.E. (S.A.)‡

(Received 15 August 1938)

[Plates 7–12]

In the course of photographic studies of lightning with the Boys' camera, which have been described in previous papers, discharges have occasionally been observed which did not seem to fit in completely with the general picture to which the majority conformed. A sufficient number of these apparently anomalous discharges has now been obtained for their peculiarities to be cleared up and for us to be able to show that their behaviour is in accord with principles identical with those which govern the more usual type of discharge.

The anomalies to be discussed are of two kinds. In the first place there are discharges which exhibit abnormal first leader processes. Some mention of these has been made in a previous paper in connexion with their electrical effects, which are characteristic and easily distinguishable (Schonland, Hodges and Collens 1938). Secondly there are discharges some of whose subsequent strokes adopt a new channel for part of their length.

1. ABNORMAL FIRST LEADER PROCESSES

In the general case the stepped leader to a first stroke travels to ground with a velocity which is approximately constant. When recorded by a camera with a slow-moving film or lens, the track of the luminous tip of the leader is a continuous line whose separation from the relatively undistorted track of the more rapid return streamer becomes uniformly less as the leader approaches the ground (Schonland, Malan and Collens 1935, fig. 2). On a faster camera the leader is resolved into a series of steps whose ends lie on a similar line.

A fast camera record of this kind is shown in the lower half of fig. 1, Plate 7: the lens of the camera was fixed and the film was moved with a

* The Bernard Price Institute of Geophysical Research, Johannesburg.

† The University of Cape Town, Associate of the Bernard Price Institute.

‡ The Victoria Falls and Transvaal Power Co., Johannesburg.

linear horizontal velocity of 50 m./sec. The time axis is in the direction of the arrow. The stepped leader starts from the base of the thundercloud at e and follows the line $eghI$ as a result of its downward movement and of the film displacement. The steps from h to I do not show on the reproduction because the very bright return portion of the stroke $IHGFE$ produced considerable halation. The leader track $eghI$ is very nearly a straight line, the leader covering the (two-dimensional) distance of 1.9 km. to the ground in 0.0077 sec. with an approximately uniform velocity of 2.5×10^7 cm./sec. Closer inspection shows that a change in velocity from 2.8×10^7 to 2.4×10^7 cm./sec. takes place after the point g , with a corresponding reduction in the length and brightness of the leader steps. It is to be noted that this is the point in the downward course of the leader at which it develops an important branch, the formation of which is shown by the horizontal stepped streak to the right of g . The change in the nature of the leader process after g is so small that in the case of this discharge the leader might well be regarded as of the normal or α type. When it occurs to a more marked degree, so as to show a pronounced discontinuous change of velocity, step-length and step-brightness at one or more points along its path, it forms the distinguishing characteristic of what has been called the type β leader (Schonland 1938). The example discussed above shows that the two types can merge into one another.

(a) *Type β_1 leaders*

The commonest type β leader (type β_1) shows a single sharp discontinuity in its downward movement. An example is given in the upper record (flash $N 2$) of fig. 1, Plate 7. The leader begins at a and travels over the first section ab of its path at an effective velocity of 7.2×10^7 cm./sec. in bright steps each about 100 m. in length. The point b on the leader, which later becomes B on the return stroke is shown by a fixed camera picture of the discharge to be the start of a long and prominent branch and the development of the leader to this branch is marked by the nearly horizontal line of steps to the right of b . From the point b onwards, the leader proceeds to ground along bcD with a considerably smaller effective velocity, 3.2×10^7 cm./sec., in steps which are so short and faint as to be discerned with difficulty on the original record.

A second example is given in fig. 2, Plate 8, which contains both the slow and the fast Boys' camera records of the first stroke of flash 43. The former is reproduced to scale in fig. 3, with an arrow to show the direction of motion of the lens.

The first part of the leader, abc , is very heavily branched and travels

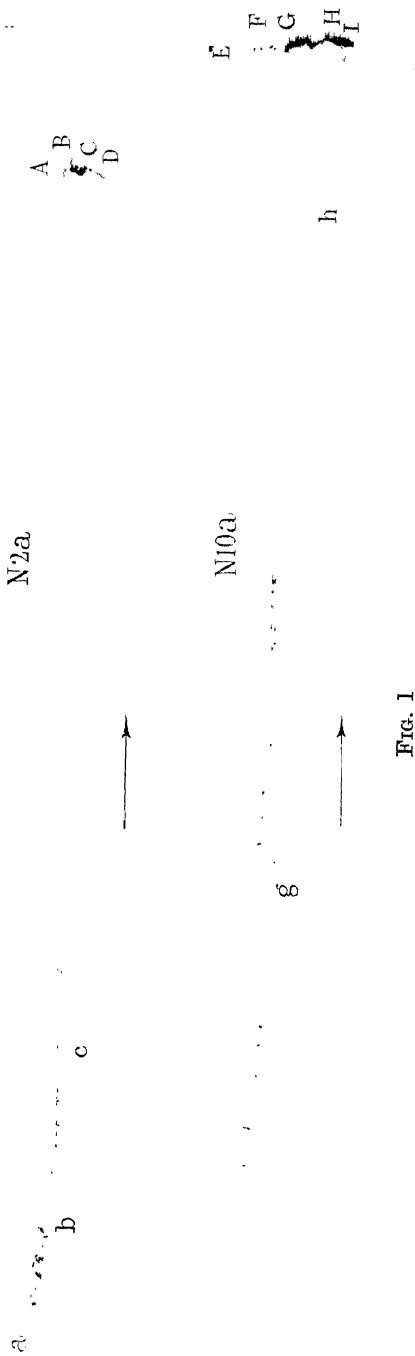


FIG. 1

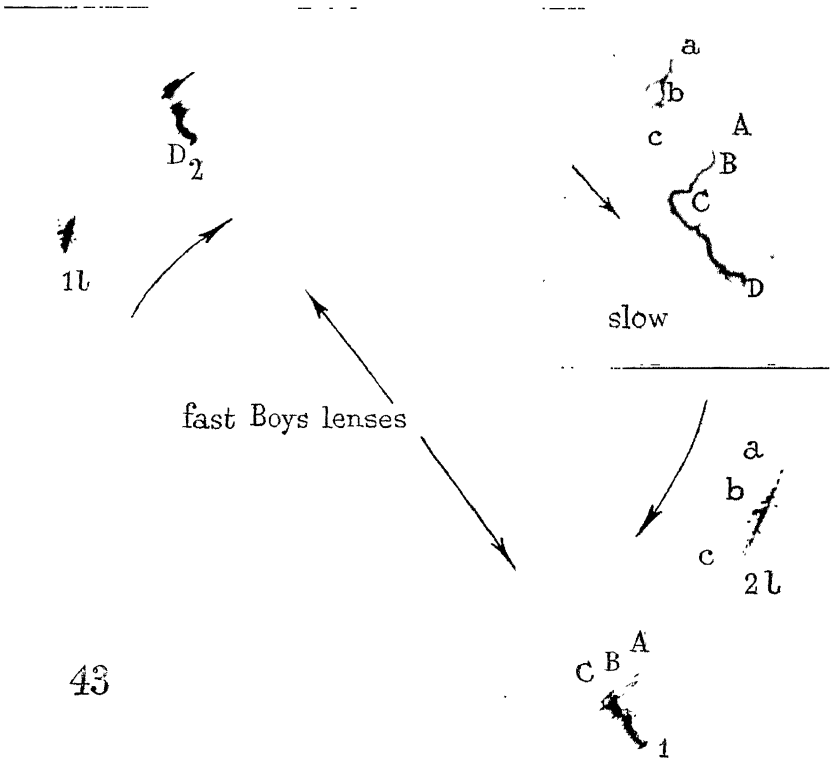


FIG. 2

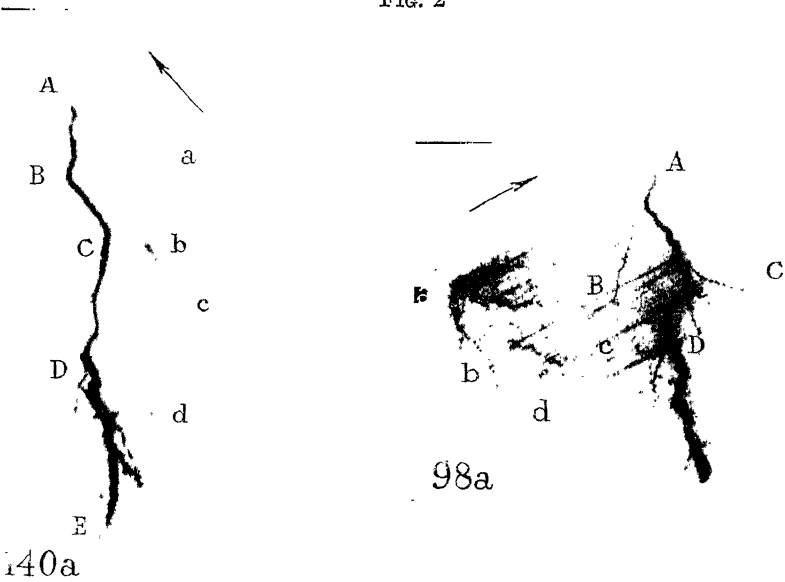


FIG. 4

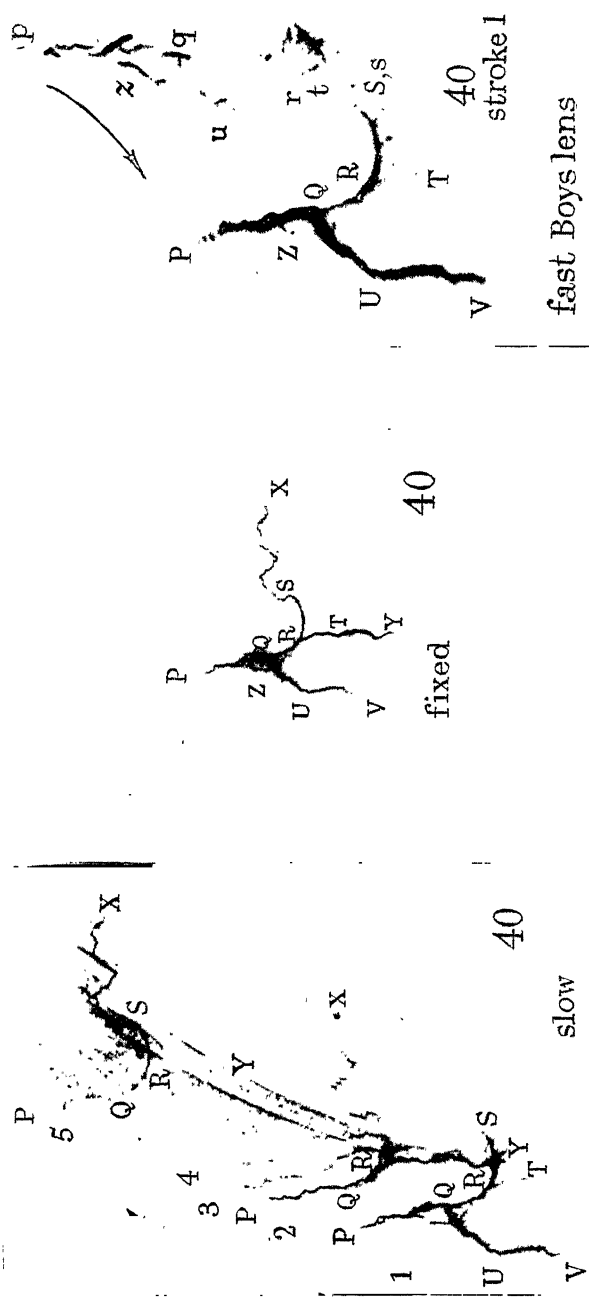


FIG. 18

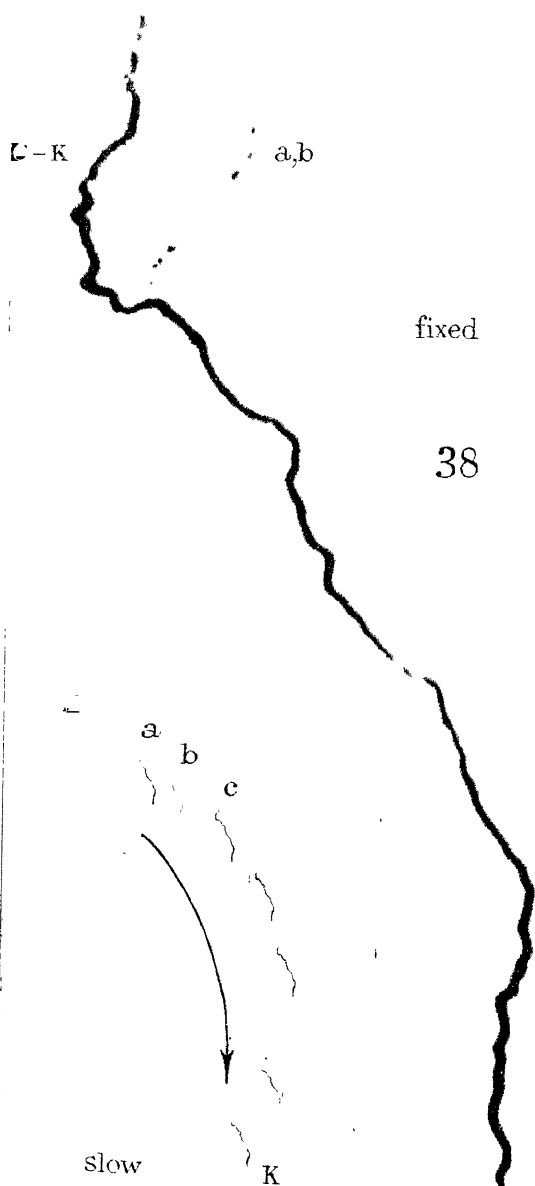


FIG. 13

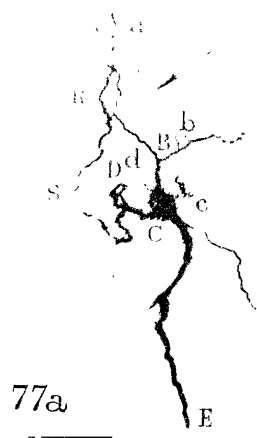


FIG. 14

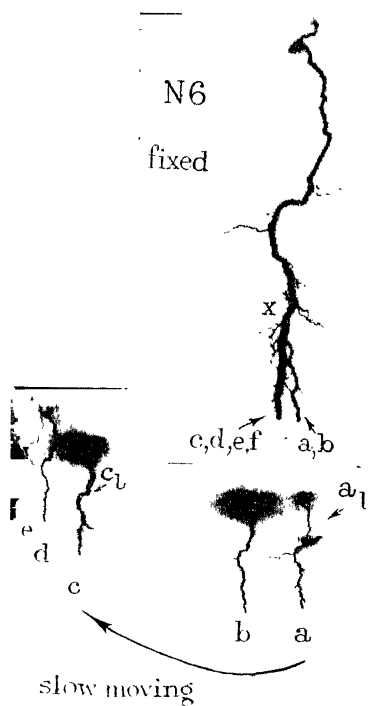


FIG. 15

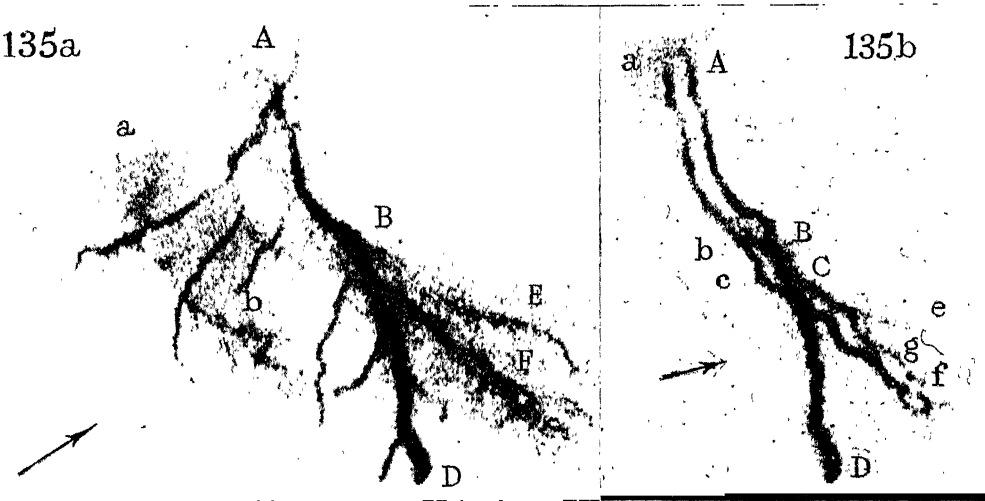


FIG. 6

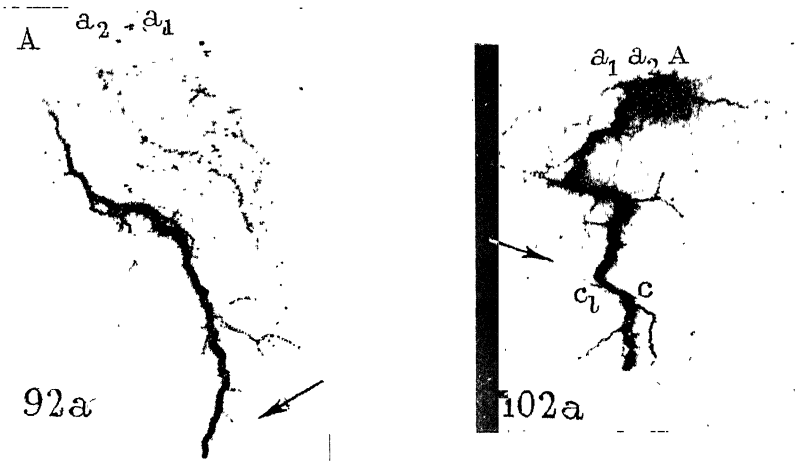


FIG. 10

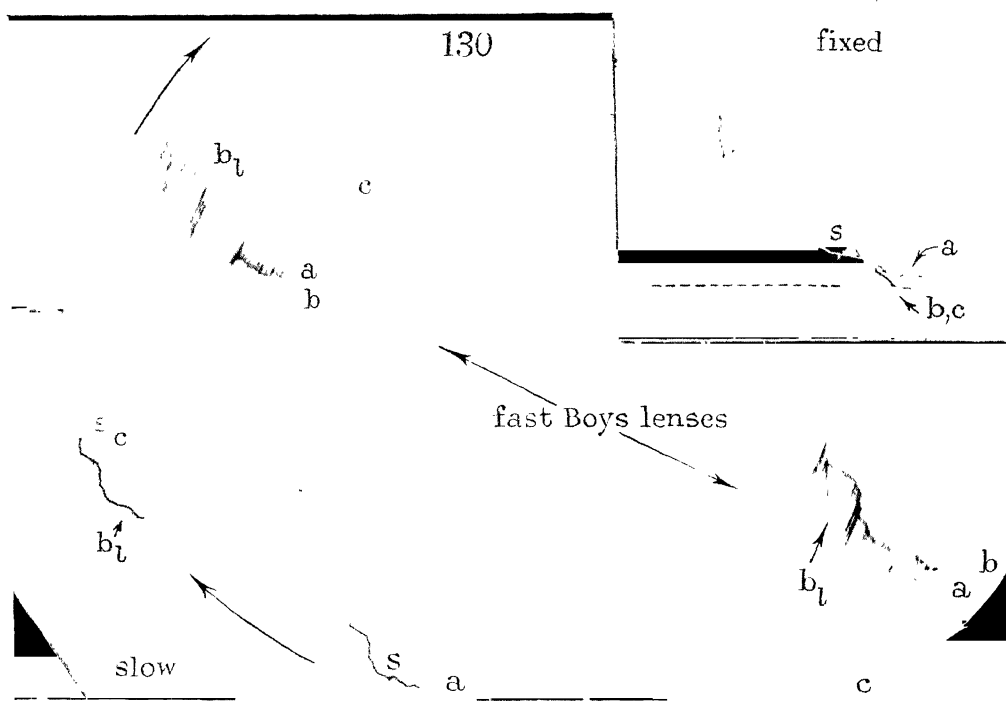


FIG. 16

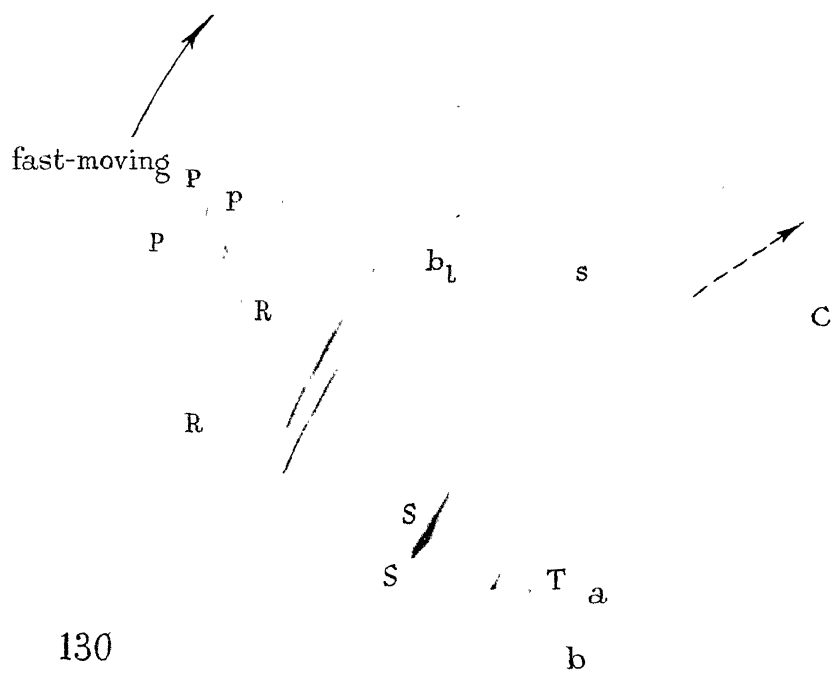


FIG. 17

from *a* to *c* at a high velocity, 1.6×10^8 cm./sec. This part of the leader is shown on the fast Boys' camera record, fig. 2, at 1*l* and 2*l*, where it is separated into a succession of very bright steps. After the leader has reached *c* the record on the slow camera is faint but can still be seen along the part shown shaded in fig. 3. The effective velocity from *c* to *D*, where the leader strikes the ground, is 1.05×10^7 cm./sec. On the fast camera this portion cannot be seen at all, since the luminosity, already weak, is stretched out by the lens motion so as to extend from the point *c* on 2*l* to *D* on the image of the return stroke marked 2 and from 1*l* to the base of that marked 1.

It will be seen from these two examples that it is a matter of some difficulty to secure a photographic record of the second and slower portion of a type β_1 leader, since the low effective velocity causes a considerable reduction in the intensity of the light emitted by the streamer. For this reason we have been able to secure very few examples of the second portion of this leader process on our fast-moving cameras and cannot always record it on the slower cameras.

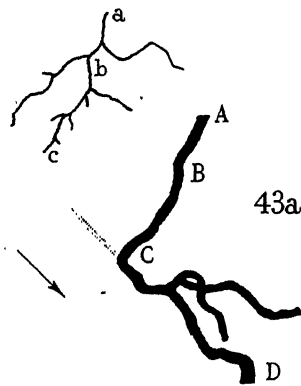


FIG. 3. Type β_1 leader; flash 43, first stroke.

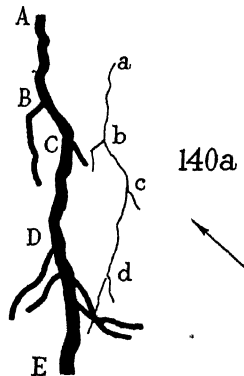


FIG. 5. Type β_1 leader; flash 140, first stroke.

Fig. 4, Plate 8 (left) shows the slow-moving camera record of the leader and return portions of stroke 140a and is reproduced to scale in fig. 5. Here the leader *abcdE* travels along the first part of its track, *abcd*, at 6.3×10^7 cm./sec., branching profusely. The final part *dE* is not recorded by the camera, but in view of the two cases already discussed and others which follow, it is presumed that the leader continues from *d* to ground with a luminosity which was too faint to be observed on this occasion. Measurement then indicates that *dE* was traversed at 1.3×10^7 cm./sec., a

velocity which is known to be so low as to require very good conditions for the streamer to be recorded photographically. Fig. 4, Plate 8 (right) shows another example, 98a, where the first portion of the leader *abcd* is bright, rapid and very heavily branched. An abrupt change to a slow and fainter portion, which is just visible on the original record, occurs in the neighbourhood of *c*. Owing to the confusion introduced by the heavy branching, measurements of the velocities cannot be made with accuracy, but the second and final portion travels at a velocity of the order of 1.0×10^7 cm./sec.

Stroke 135a is shown in fig. 6, Plate 11 (left), and reproduced to scale in fig. 7. The first part of the leader, *ab*, ends at the prominent branching point *b*, which later becomes *B* on the return portion of the stroke. The change in velocity at *b* is from more than 1.0×10^8 cm./sec. (high velocities are not easy to measure on a slow-moving camera record) to 9.0×10^6 cm./sec. The luminosity in the second stage of the leader is too weak to be recorded.

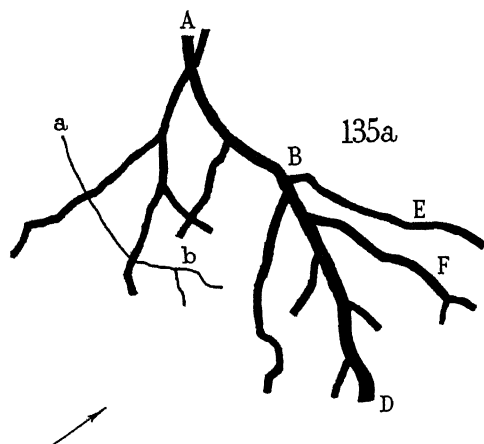


FIG. 7

FIG. 7. Type β_1 leader; flash 135, first stroke.

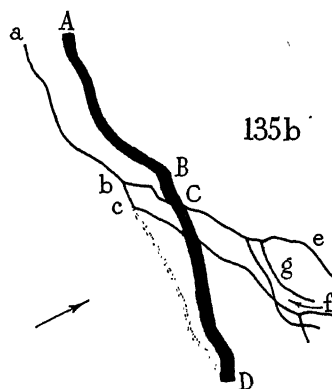


FIG. 8

FIG. 8. Flash 135; second stroke, with dart leader showing change in velocity and branching.

135b, the stroke subsequent to 135a, is interesting in connexion with the mechanism of the type β_1 leader. It is shown on the right of fig. 6, Plate 11 and in fig. 8. Since it follows the path already ionized by 135a, the leader in this case is of the high velocity dart type and brightly luminous all the way down to the ground. It is unusual, however, in showing like its predecessor an abrupt change in velocity near the point *b* and in develop-

ing strongly luminous branches at b and c . Before the point c its velocity exceeded 1.0×10^8 cm./sec., after c it fell to 5.5×10^7 cm./sec. What we consider to be a significant feature of the branching is that the return streamer $DCBA$, instead of retracing these leader branches in the usual manner, shows no sign of them whatever. Subsequent stroke leaders of this kind, though rare, have been observed a number of times and a second example is given by flash 40, which will be described later.

Table I contains the values of the effective velocities v_1 and v_2 of the two stages of type β_1 leaders in all the instances where measurement has been possible, together with the (two-dimensional) lengths of the stages and information concerning direct evidence of the second stage on the records.

TABLE I. TYPE β_1 LEADERS

Flash	First stage		Second stage		Second stage luminosity	
	Length km.	v_1 cm./sec.	Length km.	v_2 cm./sec.	Recorded	Not recorded
92	1.12	11.2×10^7	0.64	0.8×10^7	—	×
77	0.90	10.0	1.00	1.9	—	×
43	0.48	16.0	1.44	1.1	×	—
140	1.90	6.3	1.22	1.3	×	—
N2	1.00	7.2	1.54	3.2	×	—
139	1.90	10.0	1.26	1.4	—	×
40	1.15	26.0	1.90	10.5	×	—
135	1.20	> 10.0	1.20	0.9	—	×
112	0.20	—*	1.20	0.7	—	×
98	0.40	—*	0.70	(1.0)	×	—

* Not obtainable.

It will be seen that the second stage has been recorded in five out of the ten cases examined. The velocities, v_1 , associated with the first stage are all greater than 6.0×10^7 cm./sec. while those for the second stage, v_2 , with one exception, are less than 3.3×10^7 cm./sec. The exception, flash 40, which is discussed later, shows an abnormally high value of v_1 as well.

Whether the second stage is recorded by the camera or not, the measurements indicate that its velocity does not fall appreciably below 1.0×10^7 cm./sec. The values of v_1 and v_2 given in Table I are plotted, together with data from the next section, in fig. 9.

(b) Type β_2 leaders

In a rather rare variant of the type β leader, which we term type β_2 , the second and slower stage of the leader is associated with the appearance of one or more fast dart streamers, which travel rapidly down from the

cloud along the previously formed track and cease when they have caught up with the slower leader-tip. The first bright stage is thus followed after considerable intervals of time by one or more bright dart streamers, each of which travels farther than the last. Although it is difficult to record the slow and feebly luminous second stage whose progress is responsible for these later illuminations of the leader channel, it has been clearly recorded in one instance and may be presumed to have been present in all the four leaders of this type which we have observed.

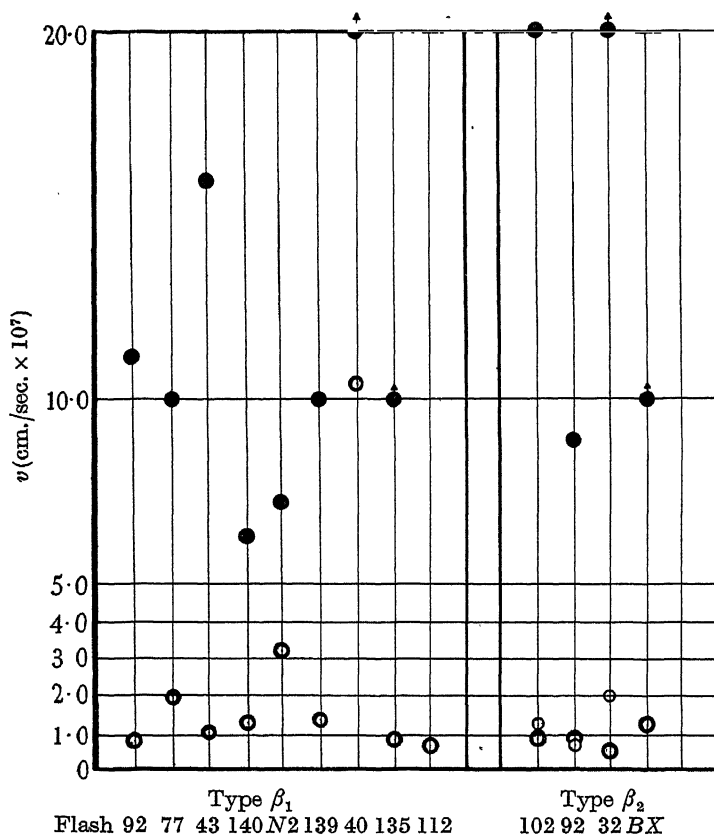


FIG. 9. Velocities in the two stages of type β leaders.
First stage, \bullet v_1 ; second stage, \circ v_2 and \circ v'_2 .

A small change in the velocity of the slow leader is usually associated with the appearance of these dart streamers.

An example of a type β_2 leader, with one dart streamer, is the first stroke of flash 102, shown in figs. 10, Plate 11, and 11. In this case the first portion of the leader, a_1b_1 (fig. 11), travels at 2.0×10^8 cm./sec. and is

bright and easily recorded. The next stage b_1c_2D , which is shown in fig. 11 as an interrupted line, is not recorded photographically. When it has reached c_2 , 0.0082 sec. after the start of the first stage from a_1 , a fast dart streamer $a_2b_2c_2$ travels along the leader channel to catch up with the slowly advancing tip at c_2 . The slow advance of the leader is then resumed along c_2D .

An increase in the effective velocity of the second stage of the leader is found to follow the appearance of this dart streamer. The velocity before c_2 , v_2 , is found by measurement to be 9.0×10^6 cm./sec. while that after it, v'_2 , is 1.3×10^7 cm./sec. The velocity of the dart streamer is greater than 2.0×10^8 cm./sec.

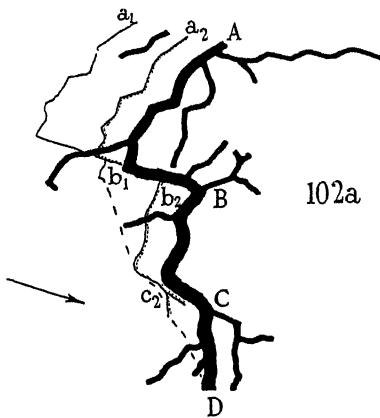


FIG. 11

FIG. 11. Type β_2 leader with dart streamer $a_2b_2c_2$; flash 102, first stroke.

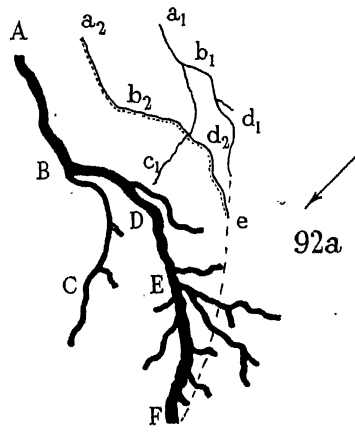


FIG. 12

FIG. 12. Type β_2 leader with dart streamer $a_2b_2d_2e$; flash 92, first stroke.

A similar behaviour is shown by flash 92, figs. 10, Plate 11, and 12. After the first stage has passed along $a_1b_1d_1$ (fig. 12) with a branch to c_1 , an apparent pause of 0.009 sec. ensues, at the close of which a bright dart streamer travels rapidly along $a_2b_2d_2$ and onwards to e . We suppose as before that during this interval a slow streamer passes from d_1 to e , to be caught up by the dart streamer at e , and continues on afterwards to reach the ground at F . The values of v_2 and v'_2 in this case are found to be 9.2×10^6 cm./sec. and 7.0×10^6 cm./sec. respectively.

Flash BX, which unfortunately cannot be reproduced, shows two such dart streamers emerging from the cloud at different times to catch up with the slower leader process in front. The interval of time between them is 0.0032 sec. and the second dart streamer travels 384 m. farther than the first. The record clearly shows the slow stepped process which continues

throughout these intervals to travel onwards to the ground. The slow process is in this instance unaffected by the occurrence of the darts and has velocities v_2 and v'_2 before and after the last dart streamer which are both 1.2×10^7 cm./sec.

Values of the three velocities associated with the three portions of the leader track in these and one other case are shown with further relevant data in table II and are plotted in fig. 9.

TABLE II. VELOCITIES AND OTHER DATA FOR TYPE β_2 LEADERS TO FIRST STROKES

	Interval before dart sec.	First stage		Second stage			
		Length km.	v_1 cm./sec.	Length km.	v_2 cm./sec.	Length km.	v'_2 cm./sec.
Flash							
102	0.0082	1.2	2.0×10^8	0.72	9.0×10^6	0.48	1.3×10^7
92	0.0090	0.5	9.0×10^7	0.80	9.2×10^6	0.64	7.0×10^6
32	0.0067	1.6	2.6×10^8	0.31	$(5.0 \times 10^6)^*$	0.61	$(2.0 \times 10^7)^*$
BX	0.0032	0.9	$> 1.0 \times 10^8$	0.40	1.2×10^7	0.40	1.2×10^7

* Values in brackets only approximate.

(c) Discussion of type β leaders

The significant features of the type β_1 leader can be summarized as follows:

(1) It begins with a rapid and brightly stepped portion which usually branches abundantly downwards from the base of the cloud. The velocity of this portion, as fig. 9 indicates, is always greater 6×10^7 cm./sec. and in one case is as high as 2.6×10^8 cm./sec.

(2) This first stage is followed, without any pause which has been detected, by a slower and much fainter one in which the leader travels down to ground with a velocity which is usually close to 1.0×10^7 cm./sec., and, except in one instance, is less than 3.3×10^7 cm./sec.

As has already been pointed out (Schonland 1938) these facts suggest that the first part of the leader is brought into being by the presence of a strong electric field between the original cloud charge and a volume charge, opposite in sign, in the air below the cloud. The heavy branching of the leader and the fact that its first portion ends at a prominent branching point are indications of the extent of the space charge, to the neutralization of which the initial portion of the leader is mainly devoted. That the whole of the space charge is not always destroyed in this way is shown by the occurrence of leaders subsequent to the first, whose branches, as in the case of stroke 135b above, repeat the branched system of the first stroke. The recoil streamers described in a previous paper (Schonland, Malan and

Collens 1935) are probably of similar origin, streamers passing from the space charge to the conducting channel of the return stroke.

The not infrequent occurrence of leaders which actually cease after executing their first and heavily branched portion (air discharges) is an indication that this portion is not controlled by an electric field reaching from the cloud all the way to the ground. Thus when the leader is found to proceed onwards along the second portion of its track we can expect it to do so at a velocity which is different from that prevailing during the earlier stage. The values given in Tables I and II and in fig. 9 show that this is the case.

It is significant that the values of v_2 (and v'_2) in fig. 9, with few exceptions, are grouped closely around 1.0×10^7 cm./sec., the minimum velocity of advance of a negative streamer into virgin air. This clearly suggests (Schonland 1938) that when space charge and the powerful electric fields associated with it are no longer playing an important part in the development of the streamer, it advances by a process of direct ionization by the electrons in its tip. If photo-ionization of the air ahead of the tip were an important factor in negative streamer development, the streamer would be moving along a previously prepared track and in low fields should show a higher velocity of advance than that observed.

What we have called the type β_2 leader does not differ in essentials from the type β_1 , since the velocity of the second stage of the leader is only slightly affected by the dart streamers which catch up with it.

The processes in a type β_2 leader appear to be very similar, though on a different scale, to those which are thought to occur (Schonland 1938) in the stepped leader itself, where a fast step streamer periodically catches up with the tip of a pilot streamer moving in front of it. Arguments have been adduced to show that in the latter case the step-interval of the order of 50μ -sec. is determined by conditions at the tip of the leader and not by processes occurring within the cloud. In the case of the steps due to dart streamers in the type β_2 leader the interval, as shown by Table II, is of the order of 0.01 sec. and we would suggest that these streamers are controlled by processes within the cloud itself, being actually new leader discharges from new centres of charge within the cloud. As in the more usual case, where however they travel all the way to ground and are the starting portions of strokes subsequent to the first, these new leaders pass down a previously prepared channel. In support of this suggestion it may be pointed out that the interval between the appearance of the first stage of a type β_2 leader and of the subsequent dart streamer, as well as the intervals between successive dart streamers when these occur, are of the same order

as the intervals between successive strokes of a normal discharge. The slowness of the leader process thus causes the type β_2 first stroke to embody in one stroke what would otherwise be two or more strokes from the cloud to ground. It would therefore be expected that type β_2 discharges would not be followed by many subsequent strokes and this seems to be borne out by the fact that three of the four discharges of Table II have no subsequent strokes while the fourth, flash 92, has only one.

A valuable illustration of the type β_2 discharge is afforded by a photograph taken by Workman, Beams and Snoddy (1936) on a camera with a slow-moving film. Their fig. 1 is described by them as a discharge which appears to pass to ground in four successive large steps. According to the discussion we have given, this would seem to have been a type β_2 discharge, the first large step corresponding to our first stepped leader stage, the next two to dart streamers superimposed on an unrecorded slow streamer in front of them and the last to the return portion of the stroke. From the data given by these authors we have calculated the velocities v_2 , v'_2 , and v''_2 of the invisible leader processes before, between and after the supposed dart streamers and have found 2.7 , 5.0 and 8.0×10^6 cm./sec. respectively. The first two values appear at first sight to be inconsistent with the view that such a process cannot travel with a velocity much below 1.0×10^7 cm./sec. but examination of the photograph shows that the portions of the track concerned in them were extremely tortuous and that the two-dimensional projection which has been used in the calculation may well be very much shorter than the actual length. The final portion, on the other hand, where the velocity is not far from that expected, is the most nearly straight of the three. The intervals between the large "steps" on this record are again close to 0.01 sec., a value which gives support to the suggestion that in a normal discharge they would be successive strokes.

2. ALTERATIONS IN THE TRACK FOLLOWED BY SUBSEQUENT STROKES OF THE SAME FLASH

The lightning discharge to ground frequently consists of a series of separate strokes each of which usually follows the same path, since its downward moving leader is guided by the ionization which has persisted since the previous stroke occurred. Instances have been described in previous papers in which, when the interval is long, this ionization is insufficient for the rapid dart leader process to continue and the leader takes a slower dart stepped form. If the ionization required to guide a leader has completely disappeared over part of the track, we may expect to observe

this part of the new leader as a stepped virgin-air process following a track different from that of its predecessors. If further strokes occur at intervals which are not abnormally long, they will follow the new channel. An effect of this kind could occur anywhere along the track of the discharge, but so far it has only been observed by us at its lower end. When such a flash is recorded on a fixed camera it appears to have forked down to ground in two or more places.

If the ionization has been destroyed by wind and lapse of time at the upper end of the channel, we might expect to find the new leader emerging from the cloud in a fresh direction and adopting the older and still ionized portion of the channel later on in its progress. On a fixed camera the flash would then appear to have branches upwards into the cloud along a Y-shaped channel. The same effect would arise if, as suggested in a previous paper (Schonland 1938), the stroke in question started from a new centre of charge within the cloud and joined up with the old channel outside instead of inside the cloud. This second explanation is to be preferred if the interval before the appearance of the new upper section of the channel is not a very long one.

Examples of these changes in the channels of subsequent strokes are given below.

(a) *Alterations at the upper end of the channel*

A flash with apparent Y-shaped upward branching is shown in the upper part of fig. 13, Plate 10 (flash 38, fixed camera picture). The slow-moving camera record below this picture indicates that the first two strokes, *a* and *b*, were preceded by leaders which came down from the right of the cloud-base while the third and following strokes, *c* to *K*, were due to leaders which emerged from a point 0.25 km. to the left of the first two. The time interval between strokes *b* and *c* was 0.065 sec., only twice the most frequent value, 0.03 sec., which has been found for the interval between strokes. Since an interval of this order does not usually cause the leader to the stroke following it to show a much reduced velocity or to take the dart-stepped form characteristic of a trail barely capable of guiding it, we suggest that the behaviour of this stroke is caused by the appearance of a leader from a centre of charge in a different part of the cloud.

In one instance only have we recorded upward branching in a single return stroke. In fig. 14, Plate 10 (flash 77, first stroke) the leader process is clearly visible along *abcd* and if it continues onwards to ground from *c* to *E* is of type β_1 . The velocity v_2 calculated for the second stage (0.8×10^7 cm./sec.) supports this suggestion. The leader also shows an

upward bend or branch from c to d and the return portion of the stroke after travelling from E to C , where the track is very contorted, branches along CBA' and $CDSRA$ to reach the cloud-base in A' and A . In this case the branching at C as far as D on the left fork, and the whole way to the cloud at A on the right fork, is in accordance with the principle that the return streamer follows the track prepared for it by the leader. The photograph, which is not a very good one, offers no clue to the mechanism of the extension of the left fork onwards from D to A' . It is, however, probable that A and A' are coincident and that two leaders travelled down from this point, the left-hand one terminating close to the fork d of the other and so providing an alternative path for the upward return process.

(b) *Alteration at the lower end of the channel*

The fixed and the slow-moving camera records of flash $N6$ are shown in fig. 15, Plate 10. On the fixed picture the channel apparently branches to the ground from the point X . The slow-moving camera shows that the first two strokes a and b took the right fork while the remaining four, from c to f , went to the left. The explanation of this effect is to be found in the unusually long time interval, 0.13 sec., between strokes b and c . This was so long that the dart leader to c was slow and can be seen as a fine line, c_l , to the right of the return portion of this stroke. It is unusual for any leader other than the slow first one, a_l , to be detected by a slow-moving camera. The velocity of c_l as far as X was 3×10^8 cm./sec., a value which for a dart leader is so low as to indicate that the track it followed was barely able to guide the leader along it. At X we must suppose that the pre-existing ionization had disappeared completely and that the leader took a new path, showing from X onwards the prolific branching which is associated with advance into virgin air. Flash 130, figs. 16, Plate 12 and 17, Plate 12, is very similar to the one just discussed. Fig. 16 shows the fixed camera record (with downward forking at S) on the right, the slow-moving camera record on the left and the double-lens Boys' camera record in the centre. Fig. 17 is an enlargement of the record of one of the Boys' camera lenses. The alteration in the channel at S was created by the leader to the second stroke, b , and can be ascribed to the exceptionally long interval of time, 0.18 sec., between the first and second strokes, as shown by the slow camera record. The leader to stroke b , which is marked b_l in figs. 16 and 17, is dart-stepped and travels at 1.7×10^8 cm./sec. as far as S . This indicates a previously ionized path which is only just able to guide the leader process. At S it becomes a stepped process advancing into virgin air along a com-

pletely new path. Both the dart-stepped and the stepped stages are clearly shown in fig. 17 where the former is visible as a broken line along b_i from p to s , the point on the leader corresponding to S on the return stroke, and the latter from s onwards as a series of disconnected spots above the dotted arrow. The steps of the dart-stepped portion have an average length of 25 m. and a pause-time of $15 \mu\text{sec.}$, while those of the stepped portion after s have the same average length and a pause-time of $65 \mu\text{sec.}$, the effective velocity changing at s from 1.7×10^8 to $3.8 \times 10^7 \text{ cm./sec.}$

In both the instances described above, the alteration in the path followed by a subsequent stroke can be directly related to the occurrence of a very long time interval between strokes and an exceptional ageing of the ionized trail. It is known that the wind sometimes causes a considerable shift in the position of this trail (Schonland 1937) and it is suggested that in these cases it has destroyed the continuity of the track at its lower end.

Flash 40, fig. 18, Plate 9, is a discharge which combines in its behaviour both the abnormalities which have been described in the present paper, for it offers a good example of a type β_1 first stroke leader and the lower portion of its second stroke takes a path which is different from that followed by the first stroke. The apparent downward forking thus produced is shown on the fixed camera picture at the centre of fig. 18.

The leader process to the first stroke is recorded in the fast Boys' lens picture on the right. The first and rapid stage of the leader begins from the cloud-base at p and travels with long and very bright steps from p to q , branching at q along qu as far as u and along qr to further branches ending at s and t . The ends of this branched first stage are the points U , S and T on the fixed picture in the centre. After reaching u the leader enters upon its second slow stage and can be seen on the original negative to travel more slowly and with reduced velocity to reach the ground at V . The branches along rs and rt also develop during this stage. The return streamer travels along $VUQP$ to the cloud and along the branch QR to S and T .

The behaviour of the strokes subsequent to this first one is indicated by the slow-moving camera record on the left of fig. 18. The second stroke, which followed the first after an interval of 0.04 sec., followed the branch qrt , which was developed by the first leader, and reached the ground at Y . In so doing it must have considerably extended its leader along qrs , for the return streamer $YT RQP$ shows the full development of this branch as far as X .

The next two strokes, (3) and (4), did not develop this branch but travelled along the new main channel to ground at Y . The last stroke of the series marked (5) in fig. 18, Plate 9, resumed the illumination of the branch

RSX in a most vigorous manner. This stroke also travelled to ground along the right-hand fork *RY*.

The behaviour of the first leader and subsequent strokes of this flash indicates the presence of a considerable space-charge in the air in the region *RSX*. The field produced by this space-charge must be considered the cause of the exceptionally high velocity (2.6×10^8 cm./sec.) of the first stage of the first leader and of its exceedingly bright steps. It is probable that during the interval between first and second strokes the ionization along the path *QUV* disappeared completely. The second stroke was then guided by the space-charge to ground along *TY* and into the air along *RSX*. The next two strokes occurred at intervals of 0.010 and 0.015 sec. after their predecessors, and followed the same channel to ground. The last stroke took place after a longer interval, 0.063 sec., and presumably was again influenced by the space-charge to develop strongly in the direction *RSX*. The record shows that some form of partial discharge occurred between the fourth and last strokes, the nature of which cannot be properly determined.

This work is part of the programme of lightning research sponsored by the South African Institution of Electrical Engineers. Our thanks are due to Mr J. A. Linton for his help in constructing the new camera used in obtaining fig. 1, Plate 7 and other data.

SUMMARY

An account is given of some apparently anomalous lightning discharges observed with the Boys' camera, and it is shown that their behaviour is in accordance with the same principles as govern the more usual type of discharge.

Certain abnormal first leader processes occur in two stages, the first rapid and the second involving a velocity close to the limiting value for negative streamer propagation. The first stage is considered to arise from the presence of space-charge concentrations in the air below the cloud. Dart streamers may travel down from the cloud during the second slow stage and catch up with the leader tip.

Changes in the lower portions of the channels used by successive strokes of the same flash are shown to be associated with long time intervals between strokes, whereby the conductivity and guiding power of the previous track is destroyed. A case of change in the upper portion of the

channel is ascribed to the existence of a second centre of charge within the cloud.

REFERENCES

- Schonland, B. F. J. 1937 *Phil. Mag.* **23**, 503.
— 1938 *Proc. Roy. Soc. A*, **164**, 132.
Schonland, B. F. J., Hodges, D. B. and Collens, H. 1938 *Proc. Roy. Soc. A*, **166**, 56.
Schonland, B. F. J., Malan, D. J. and Collens, H. 1935 *Proc. Roy. Soc. A*, **152**, 595.
Workman, E. J., Beams, J. W. and Snoddy, L. B. 1936 *Physics*, **7**, 375.
-

A new Coriolis perturbation in the methane spectrum*

I. Vibrational-rotational Hamiltonian and
wave functions

BY H. A. JAHN

*Davy-Faraday Laboratory, The Royal Institution**(Communicated by Sir William Bragg, O.M., P.R.S.—**Received 25 May 1938)*

INTRODUCTION

It is known that the infra-red absorption band of the low-frequency fundamental vibration (ν_4) of the methane molecule has a rotational structure which is much more complex than one would expect for such a simple molecule (Nielsen and Nielsen 1935). It is our purpose to show that nevertheless this complex structure can be explained on the basis of a regular

* The term Coriolis perturbation is used to cover all perturbations in polyatomic molecules which arise from the interaction of the angular momentum of degenerate vibrations with the rotational angular momentum of the molecule (cf. Teller (1934) or Johnston and Dennison (1935)). The additional forces which arise in a rotating mechanical system were investigated by Coriolis (*Journal Polytechnique*, 1832); see, for instance, Routh (1898), *Dynamics of a Particle*, p. 156 or Haas (1928), *Theoretical Physics*, I, 45. The gyroscopic nature of the Coriolis (or compound centripetal) acceleration was pointed out by Thomson and Tait (1879), *Natural Philosophy*, I, 392: see Webster (1912), *Dynamics*, p. 320.

tetrahedral model for the molecule. We shall see that the rotational levels of the vibration ν_4 are perturbed by the rotational levels of the next nearest vibration ν_2 in such a way as to produce in the spectrum just the observed complex structure. The perturbation arises from a Coriolis (or vibrational gyroscopic) interaction between the rotational-vibrational levels of the two different modes of vibration. In this first part we derive these Coriolis coupling terms in the vibrational-rotational Hamiltonian and find also the correct rotational-vibrational wave functions with which to carry out the perturbation calculation. In Part II we evaluate the matrix elements of the perturbation and determine the theoretical rotational energy spectrum of ν_4 . In the third and final part we use this energy spectrum to calculate the optical spectrum and compare this with the observed spectrum. For this purpose we calculate the theoretical intensities of the rotational fine structure lines, taking into account the nuclear spin weights of the four equivalent hydrogen atoms. From this theoretical spectrum we calculate the theoretical envelope which would be observed with slit widths of approximately 0.5 cm^{-1} as used by Nielsen and Nielsen. This theoretical envelope is found to agree remarkably well with the experimental envelope, even without taking into account any vibrational or rotational change in the equilibrium configuration.

1. PURE DEFORMATION AND ORTHOGONAL VALENCY MODES

Before we can make any explicit calculation of the rotational-vibrational levels of the methane molecule we must first of all know the fundamental modes of vibration. In the following we make use of what we call the pure deformation and orthogonal valency modes (cf. Mecke 1930). These are not exact normal modes of vibration of the molecule, but we shall find that they are very good approximations to the true modes. They are based on the experimental fact that the energy required to stretch a C—H bond is considerably greater than the energy required to change the H—C—H angles. Thus we can obtain a good approximation to the low-frequency modes by introducing the condition that all the C—H distances remain invariant. This removes four degrees of freedom and enables us to find five different orthogonal deformation vibrations in which only the angles change. The remaining four high-frequency modes are then determined simply by the condition that they should be orthogonal to these deformation vibrations. These high frequency valency vibrations will involve essentially changes in the C—H distances; they will also, however, involve to a slight extent

changes in the angles. In finding these modes it will help us considerably to make use of the fact that three normal modes of vibration of the methane molecule, for any force system whatever (consistent with the tetrahedral symmetry), are determined completely by symmetry. These are the totally symmetrical vibration of type A_1 (in the notation of Tisza 1933) and the twofold degenerate set of vibrations of type E. That these are completely determined is a consequence of the group-theoretical result that the methane molecule possesses one and only one of each of these irreducible types of vibration. We denote the vibration of type A_1 by Q_1 and two suitably chosen normal modes of type E by Q_{2a} and Q_{2b} . In each of these modes of vibration the carbon atom remains at rest, in Q_1 the hydrogen atoms move radially in phase either away or towards the carbon atom (so that this is a pure valency mode of vibration), whilst in Q_{2a} and Q_{2b} the hydrogen atoms move on a sphere about the carbon atom, so that the C—H distances remain unaltered. These latter give us therefore at once two of the deformation modes of vibration. In the three remaining deformation modes the carbon atom takes part in the vibration. These modes can be arrived at by regarding them as compounded out of an infinitesimal translation of the whole molecule, followed by a backward displacement of the hydrogen atoms tangential to the equilibrium sphere drawn with its centre at the displaced position of the carbon atom. This will clearly satisfy the condition that the C—H distances remain unchanged, since neither of the two compounded displacements change them. The directions of displacement of the hydrogen atoms tangential to the sphere are determined completely by the conditions of orthogonality to the translations, the rotations and the twofold degenerate vibrations. In the three remaining valency modes the carbon atom also moves and the hydrogen atoms move out or in radially towards the equilibrium position of the central atom (and not towards the carbon atom itself) so that the valency angles do change slightly.

In the above way we found the following set of valency and deformation modes of vibration of the methane molecule. We use here the notation used previously (Jahn 1935): each mode of vibration Q_r is described by a set of vectors $u_{\alpha r}$ ($\alpha = 0, \dots, 4$),

$$Q_r = (u_{0r}; u_{1r}, u_{2r}, u_{3r}, u_{4r}),$$

giving the displacements from equilibrium of the five atoms. (The index 0 refers to the carbon atom, and the four hydrogen atoms are numbered as in the reference just mentioned.) These displacement vectors are expressed as linear combinations of the unit vectors x, y, z drawn parallel to the two-fold axes of the equilibrium tetrahedron.

(a) *Deformation modes* (CH distances do not change).

$$E \sim \begin{cases} Q_{2a} = \frac{1}{2\sqrt{2}} (0; x-y, -x+y, -x-y, x+y), \\ Q_{2b} = \frac{1}{2\sqrt{6}} (0; x+y-2z, -x-y-2z, -x+y+2z, x-y+2z); \end{cases}$$

$$F_2 \sim \begin{cases} Q_{4x} = \frac{\sqrt{M}}{2\sqrt{\{2(2m_0+M)\}}} \\ \quad \times \left(-\frac{8m}{M}x; \frac{2m_0}{M}x-y-z, \frac{2m_0}{M}x-y+z, \frac{2m_0}{M}x+y-z, \frac{2m_0}{M}x+y+z \right), \\ Q_{4y} = \frac{\sqrt{M}}{2\sqrt{\{2(2m_0+M)\}}} \\ \quad \times \left(-\frac{8m}{M}y; \frac{2m_0}{M}y-x-z, \frac{2m_0}{M}y-x+z, \frac{2m_0}{M}y+x+z, \frac{2m_0}{M}y+x-z \right), \\ Q_{4z} = \frac{\sqrt{M}}{2\sqrt{\{2(2m_0+M)\}}} \\ \quad \times \left(-\frac{8m}{M}z; \frac{2m_0}{M}z-x-y, \frac{2m_0}{M}z+x+y, \frac{2m_0}{M}z-x+y, \frac{2m_0}{M}z+x-y \right). \end{cases}$$

(b) *Valency modes* (orthogonal to deformation modes).

$$A_1 \sim Q_1 = \frac{1}{2\sqrt{3}} (0; x+y+z, -x-y+z, -x+y-z, x-y-z);$$

$$F_2 \sim \begin{cases} Q_{3x} = \frac{\sqrt{m_0}}{2\sqrt{(2m_0+M)}} \\ \quad \times \left(-\frac{4m}{m_0}x; x+y+z, x+y-z, x-y+z, x-y-z \right), \\ Q_{3y} = \frac{\sqrt{m_0}}{2\sqrt{(2m_0+M)}} \\ \quad \times \left(-\frac{4m}{m_0}y; x+y+z, x+y-z, -x+y-z, -x+y+z \right), \\ Q_{3z} = \frac{\sqrt{m_0}}{2\sqrt{(2m_0+M)}} \\ \quad \times \left(-\frac{4m}{m_0}z; x+y+z, -x-y+z, x-y+z, -x+y+z \right). \end{cases}$$

Here m = mass of hydrogen atom,
 m_0 = mass of carbon atom,
 $M = m_0 + 4m$ = total mass of molecule.

The normalization factors have been so chosen that the reduced masses of the vibrations are all equal to m , the mass of the hydrogen atom:

$$\mu_r = \sum_{\alpha} m_{\alpha} (\mathbf{u}_{\alpha r}, \mathbf{u}_{\alpha r}) = m.$$

Hence if we introduce coordinates q_r to describe a general displaced configuration Q by

$$Q = Q_0 + \sum_r q_r Q_r,$$

where Q_0 is the equilibrium configuration, then the vibrational kinetic energy has the form

$$2T = m \sum_r \dot{q}_r^2.$$

From group theory it follows, since the sets of displacements Q_{3x}, Q_{3y}, Q_{3z} and Q_{4x}, Q_{4y}, Q_{4z} are both transformed in the same way according to the representation F_2 of the tetrahedral group by the Wigner symmetry operators, that the potential energy has the following form:

$$2V = k_1 q_1^2 + k_2 (q_{2a}^2 + q_{2b}^2) + k_3 (q_{3x}^2 + q_{3y}^2 + q_{3z}^2) \\ + k_4 (q_{4x}^2 + q_{4y}^2 + q_{4z}^2) + 2k_{34} (q_{3x} q_{4x} + q_{3y} q_{4y} + q_{3z} q_{4z}).$$

If the above set of displacements were true normal modes, k_{34} would be zero. This is not the case, but it can be shown that it is small compared with the other constants.*

That these displacements do form a good approximation to the true modes can be shown further by calculating the maximum angular momenta (ζ -values) of the two threefold degenerate modes of vibration. It has been

* [Note added in proof.] The positive value of Rosenthal's general force constant D (*Phys. Rev.* 45, 538 (1934)) corresponds to a small value $k_{34} = -0.1522 \times 10^5$ of the cross term, whilst the negative value corresponds to the large value $k_{34} = -0.9957 \times 10^5$. Dennison and Johnston (*Phys. Rev.* 47, 93 (1935)) have shown that the positive value of D is required to explain the spectrum of CH_3D , thus giving conclusive evidence that the valency and deformation modes are good approximations. For the true modes, using the small value of k_{34} , we find

$$Q'_3 = 0.9994Q_3 - 0.0346Q_4, \\ Q'_4 = 0.0346Q_3 + 0.9994Q_4.$$

It should be pointed out that a positive value of Rosenthal's constant D corresponds to a negative value of Johnson and Dennison's (*Phys. Rev.* 48, 868 (1935)) constant e .

shown (Jahn 1935; cf. also Eckart 1935) that the three components S_j ($j = x, y, z$) of the angular momentum of the vibrations are given by

$$\begin{aligned} S_j &= \frac{\hbar}{i} \sum_{s>r} \frac{C_j^{(rs)}}{\mu_s} \left(q_r \frac{\partial}{\partial q_s} - q_s \frac{\partial}{\partial q_r} \right) \\ &= \sum_{s>r} \frac{C_j^{(rs)}}{\mu_s} (q_r p_s - q_s p_r), \end{aligned}$$

where the coefficients $C_j^{(rs)}$ or $C_x^{(rs)}$, $C_y^{(rs)}$, $C_z^{(rs)}$ are the scalar products

$$\begin{aligned} C_x^{(rs)} &= \sum_{\alpha} m_{\alpha} (\mathbf{x} \mathbf{u}_{\alpha r}, \mathbf{u}_{\alpha s}) = \{x Q_r, Q_s\}, \\ C_y^{(rs)} &= \{y Q_r, Q_s\}, \\ C_z^{(rs)} &= \{z Q_r, Q_s\}, \end{aligned}$$

$x Q_r$ being the set of displacement vectors obtained from Q_r by replacing each vector $\mathbf{u}_{\alpha r}$ of Q_r by the vector product $\mathbf{x} \mathbf{u}_{\alpha r}$ of the unit vector \mathbf{x} with $\mathbf{u}_{\alpha r}$ and similarly for $y Q_r$ and $z Q_r$. Comparison with Teller's (1934) definition of the ζ -values of the threefold degenerate modes of vibration shows that

$$\zeta_3 = \frac{\{x Q_{3y}, Q_{3z}\}}{m}, \quad \zeta_4 = \frac{\{x Q_{4y}, Q_{4z}\}}{m}.$$

The simple calculation gives at once

$$\zeta_3 = \frac{4m}{3m_0 + 4m}, \quad \zeta_4 = \frac{3m_0 - 4m}{2(3m_0 + 4m)},$$

in agreement with Teller's result that the sum of the ζ -values must equal $\frac{1}{2}$ independent of the force system:

$$\zeta_3 + \zeta_4 = \frac{3m_0 + 4m}{2(3m_0 + 4m)} = \frac{1}{2}.$$

Substituting the values for the masses, we find

$$\zeta_3 = 0.10, \quad \zeta_4 = 0.40.$$

The experimental values are (Childs 1936)

$$\zeta_3 = 0.05, \quad \zeta_4 = 0.45.$$

This shows that the modes we use do approximate to the true modes of vibration of the molecule. (It should be noted that the formulae given above allow us to calculate the ζ -values for other tetrahedral molecules for which valency and deformation modes are also good approximations. Under this assumption we find for GeH_4 $\zeta_3 = 0.018$, $\zeta_4 = 0.482$ and for SiH_4 $\zeta_3 = 0.046$, $\zeta_4 = 0.454$.)

2. THE CORIOLIS COUPLING OF ν_4 AND ν_2

With the above explicit expressions for the modes of vibration we can calculate, using a method already developed (Jahn 1935), all the Coriolis coupling terms occurring in the rotational-vibrational Hamiltonian of the molecule. We want in particular to calculate the Coriolis coupling between the fundamentals ν_4 and ν_2 .

The Coriolis coupling term in question can be written in the form

$$H^{(24)} = - \frac{L_x S_x^{(24)} + L_y S_y^{(24)} + L_z S_z^{(24)}}{A},$$

where A is the moment of inertia, L_x , L_y , L_z are the components of the total angular momentum of the molecule and $S_x^{(24)}$, $S_y^{(24)}$, $S_z^{(24)}$ are components of vibrational angular momentum obtained by combining the displacements of ν_4 with the momenta of ν_2 and vice versa. (Cf. Teller 1934, who considers, however, only the coupling of vibrations of the same symmetry type.) These components of angular momentum are included in the general expression given above for the angular momentum of the vibrations. The only coefficients $C_j^{(24)}$ which do not vanish are the following

$$C_x^{(2a, 4x)} = \frac{\sqrt{M}}{2\sqrt{(2m_0 + M)}} m = C_y^{(2a, 4y)},$$

$$C_z^{(2a, 4z)} = - \frac{\sqrt{M}}{\sqrt{(2m_0 + M)}} m,$$

$$C_x^{(2b, 4x)} = - \frac{\sqrt{3}\sqrt{M}}{2\sqrt{(2m_0 + M)}} m = - C_y^{(2b, 4y)}.$$

We thus find the following expressions for the components of vibrational angular momentum occurring in $H^{(24)}$:

$$S_x^{(24)} = \zeta_{24} \frac{\hbar}{i} \left(\frac{1}{2} \left(q_{2a} \frac{\partial}{\partial q_{4x}} - q_{4x} \frac{\partial}{\partial q_{2a}} \right) - \frac{\sqrt{3}}{2} \left(q_{2b} \frac{\partial}{\partial q_{4x}} - q_{4x} \frac{\partial}{\partial q_{2b}} \right) \right),$$

$$S_y^{(24)} = \zeta_{24} \frac{\hbar}{i} \left(\frac{1}{2} \left(q_{2a} \frac{\partial}{\partial q_{4y}} - q_{4y} \frac{\partial}{\partial q_{2a}} \right) + \frac{\sqrt{3}}{2} \left(q_{2b} \frac{\partial}{\partial q_{4y}} - q_{4y} \frac{\partial}{\partial q_{2b}} \right) \right),$$

$$S_z^{(24)} = - \zeta_{24} \frac{\hbar}{i} \left(q_{2a} \frac{\partial}{\partial q_{4z}} - q_{4z} \frac{\partial}{\partial q_{2a}} \right),$$

with

$$\zeta_{24} = \sqrt{\frac{M}{2m_0 + M}}.$$

The Coriolis coupling term $H^{(24)}$ is an operator which acting on the vibrational-rotational wave functions of ν_4 converts them into vibrational-rotational wave functions of ν_2 and vice versa. It can thus produce a perturbation in the rotational levels of ν_4 and ν_2 . Now Teller's Coriolis interaction term

$$H^{(4)} = - \frac{L_x S_x^{(4)} + L_y S_y^{(4)} + L_z S_z^{(4)}}{A}$$

of the angular momentum $S^{(4)}$ of the vibration ν_4 alone with the total angular momentum of the molecule has spherical symmetry and consequently does not completely remove the degeneracy of the rotational levels of ν_4 . As is well known it causes each rotational level (except the $J = 0$ level) to split into three Coriolis levels (ζ -levels). The term $H^{(24)}$, however, has only tetrahedral symmetry and consequently can partly remove the degeneracy of these Coriolis levels. This causes a splitting as well as a displacement of the rotational levels and gives rise to the observed complex rotational structure of ν_4 .

Since the term $H^{(24)}$ is totally symmetrical with respect to the tetrahedral group T_d (as can easily be verified), it can only have non-vanishing matrix elements between wave functions which transform according to the same irreducible representations of T_d . It is our purpose in the following to find those correct linear combinations of the vibrational-rotational wave functions of ν_4 which transform according to such irreducible representations. We shall carry out this explicit calculation up to the tenth rotational quantum number.

3. WAVE FUNCTIONS FOR TELLER'S CORIOLIS LEVELS

We denote the vibrational wave functions of ν_4 which correspond to one quantum of vibration in the configurations Q_{4x} , Q_{4y} , Q_{4z} by v_x , v_y , v_z . It will be convenient to introduce the linear combinations

$$v_1 = -\frac{v_x - iv_y}{\sqrt{2}}, \quad v_0 = v_z, \quad v_{-1} = \frac{v_x + iv_y}{\sqrt{2}}.$$

These functions transform according to the odd representation D_1^u of the group D_∞^i of all rotations and reflexions in the space of the threefold degenerate vibrations. The rotational wave functions w_{MK}^J of the spherical top transform, for even and odd values of J respectively, according to the representations D_0^g , D_0^u of the group D_∞^i of all rotations and reflexions about axes fixed in the molecule (cf. Jahn 1935). Consequently the vibrational-

rotational wave functions transform according to the product representations

$$D_1^u \times D_J^q = D_{J-1}^u + D_J^u + D_{J+1}^u \quad (J \text{ even}),$$

$$D_1^u \times D_J^u = D_{J-1}^q + D_J^q + D_{J+1}^q \quad (J \text{ odd}).$$

The reduction of these product representations gives us the wave functions of the three Coriolis levels into which each rotational level is split by the Coriolis interaction term $H^{(4)}$. We denote the corresponding levels by J_{J-1} , J_J , J_{J+1} respectively. The formulae giving the explicit reduction of these product representations are well known and can be written as follows:

$$\begin{aligned} W_K(J_{J-1}) = & \sqrt{\frac{(J+K)(J+K+1)}{2J(2J+1)}} w_{K+1}^J v_{-1} - \sqrt{\frac{(J+K)(J-K)}{J(2J+1)}} w_K^J v_0 \\ & + \sqrt{\frac{(J-K)(J-K+1)}{2J(2J+1)}} w_{K-1}^J v_1 \quad (K = J-1, \dots, -J+1), \end{aligned}$$

$$\begin{aligned} W_K(J_J) = & \sqrt{\frac{(J-K)(J+K+1)}{2J(J+1)}} w_{K+1}^J v_{-1} + \frac{K}{\sqrt{J(J+1)}} w_K^J v_0 \\ & - \sqrt{\frac{(J+K)(J-K+1)}{2J(J+1)}} w_{K-1}^J v_1 \quad (K = J, \dots, -J), \end{aligned}$$

$$\begin{aligned} W_K(J_{J+1}) = & \sqrt{\frac{(J-K)(J-K+1)}{2(J+1)(2J+1)}} w_{K+1}^J v_{-1} + \sqrt{\frac{(J+K+1)(J-K+1)}{(J+1)(2J+1)}} w_K^J v_0 \\ & + \sqrt{\frac{(J+K+1)(J+K)}{2(J+1)(2J+1)}} w_{K-1}^J v_1 \quad (K = J+1, \dots, -J-1). \end{aligned}$$

(We have omitted the magnetic quantum number M , since it does not enter into the transformation formulae; in particular we could take $M = 0$ for which case the rotational wave functions become simply spherical harmonics. None of the Coriolis interaction terms remove the $(2J+1)$ -fold degeneracy which is described by the quantum number M .) These vibrational-rotational wave functions diagonalize the interaction term $H^{(4)}$. One finds easily for the matrix elements (cf. Johnston and Dennison 1935):

$$\left(J_{J-1} \left| \frac{(L-S^{(4)})^2}{2A} \right| J_{J-1} \right) = \frac{\hbar^2}{2A} \{ J(J+1) - 2J\zeta_4 + 2\zeta_4(\zeta_4-1) \},$$

$$\left(J_J \left| \frac{(L-S^{(4)})^2}{2A} \right| J_J \right) = \frac{\hbar^2}{2A} \{ J(J+1) + 2\zeta_4(\zeta_4-1) \},$$

$$\left(J_{J+1} \left| \frac{(L-S^{(4)})^2}{2A} \right| J_{J+1} \right) = \frac{\hbar^2}{2A} \{ J(J+1) + 2(J+1)\zeta_4 + 2\zeta_4(\zeta_4-1) \}.$$

Since these Coriolis levels have spherical symmetry (i.e. transform according to irreducible representations of the full group of all rotations and reflexions), the number and degeneracy of the sublevels into which they can be split by the tetrahedrally symmetrical perturbation $H^{(24)}$ are determined by the reduction of the representations D_J^y , D_J^z into irreducible representations of the tetrahedral group. The general result of this reduction is easy to calculate and is given below, together with an explicit reduction carried out up to $J = 10$.

4. TETRAHEDRALLY IRREDUCIBLE HARMONIC FUNCTIONS

The reduction of the representations D_J^y , D_J^z into irreducible representations of the tetrahedral group T_d is determined by the characters of the former representations for the group elements of T_d . These characters are as follows:

Class	E	$8C_3$	$3C_2$	$6\sigma_d$	$6S_4$
ϕ	0	$\pm \frac{2}{3}\pi$	π	π	$\mp \frac{1}{2}\pi$
D_J^y	$2J+1$	$\frac{\sin(J+\frac{1}{2})\frac{2}{3}\pi}{\sin\frac{1}{3}\pi}$	$\frac{\sin(J+\frac{1}{2})\pi}{\sin\frac{1}{2}\pi}$	$\frac{\sin(J+\frac{1}{2})\pi}{\sin\frac{1}{2}\pi}$	$\frac{\sin(J+\frac{1}{2})\frac{1}{2}\pi}{\sin\frac{1}{4}\pi}$
D_J^z	$2J+1$	$\frac{\sin(J+\frac{1}{2})\frac{2}{3}\pi}{\sin\frac{1}{3}\pi}$	$\frac{\sin(J+\frac{1}{2})\pi}{\sin\frac{1}{2}\pi}$	$-\frac{\sin(J+\frac{1}{2})\pi}{\sin\frac{1}{2}\pi}$	$-\frac{\sin(J+\frac{1}{2})\frac{1}{2}\pi}{\sin\frac{1}{4}\pi}$

From these characters and from the characters of the irreducible representations of T_d one finds the reductions given in Table I. General formulae for any value of J can easily be found but are not necessary for our purpose. Our results include those of Cundy (1938) who has given the reduction of D_J^y for J even and of D_J^z for J odd.

To carry out the explicit reduction we make use of the spherical harmonics Y_K^J ($K = J, \dots, -J$), which transform according to D_J^y for J even and according to D_J^z for J odd. Having reduced the representations subtended by these functions, it is straightforward to obtain the reduction in the remaining cases (i.e. D_J^y for J odd and D_J^z for J even). We follow Maxwell (1892) (cf. Elert 1928; Bethe 1929) and express the spherical harmonics as multipoles having their axes along the x , y and z axes. Their transformation properties with respect to the tetrahedral group are then easy to investigate.

We replace the usual complex spherical harmonics Y_K^J (as defined for

instance by Bethe 1933) for each value of J by the following $2J+1$ real and orthogonal functions:

$$\left. \begin{aligned} U_K^J &= \frac{1}{\sqrt{2}} \{ Y_{-K}^J + (-1)^K Y_K^J \}, \\ V_K^J &= \frac{1}{i\sqrt{2}} \{ Y_{-K}^J - (-1)^K Y_K^J \}, \end{aligned} \right\} \quad (K = 1, \dots, J).$$

$$W^J = Y_0^J.$$

TABLE I

J	Reduction of D_0^J	Reduction of D_0^J
0	A_1	A_2
1	F_1	F_2
2	$E + F_2$	$E + F_1$
3	$A_2 + F_1 + F_2$	$A_1 + F_1 + F_2$
4	$A_1 + E + F_1 + F_2$	$A_2 + E + F_1 + F_2$
5	$E + 2F_1 + F_2$	$E + F_1 + 2F_2$
6	$A_1 + A_2 + E + F_1 + 2F_2$	$A_1 + A_2 + E + 2F_1 + F_2$
7	$A_2 + E + 2F_1 + 2F_2$	$A_1 + E + 2F_1 + 2F_2$
8	$A_1 + 2E + 2F_1 + 2F_2$	$A_2 + 2E + 2F_1 + 2F_2$
9	$A_1 + A_2 + E + 3F_1 + 2F_2$	$A_1 + A_2 + E + 2F_1 + 3F_2$
10	$A_1 + A_2 + 2E + 2F_1 + 3F_2$	$A_1 + A_2 + 2E + 3F_1 + 2F_2$
11	$A_2 + 2E + 3F_1 + 3F_2$	$A_1 + 2E + 3F_1 + 3F_2$
12	$2A_1 + A_2 + 2E + 3F_1 + 3F_2$	$A_1 + 2A_2 + 2E + 3F_1 + 3F_2$
13	$A_1 + A_2 + 2E + 4F_1 + 3F_2$	$A_1 + A_2 + 2E + 3F_1 + 4F_2$
14	$A_1 + A_2 + 3E + 3F_1 + 4F_2$	$A_1 + A_2 + 3E + 4F_1 + 3F_2$
15	$A_1 + 2A_2 + 2E + 4F_1 + 4F_2$	$2A_1 + A_2 + 2E + 4F_1 + 4F_2$

From Maxwell's discussion (*loc. cit.*), we find the following expressions for these functions in terms of multipoles having their axes along the x , y and z directions:

$$U_K^J = \frac{\sqrt{2} J!}{\sqrt{\{(J+K)!(J-K)!\}}} a_K^{(J-K)},$$

$$V_K^J = \frac{\sqrt{2} J!}{\sqrt{\{(J+K)!(J-K)!\}}} b_K^{(J-K)},$$

$$W^J = (0, 0, J),$$

where $(0, 0, J)$ is the normalized multipole directed along the z axis defined by

$$(0, 0, J) = (-1)^J \frac{\sqrt{(2J+1)}}{J!} \left\{ \frac{d^J}{dz^J} \left(\frac{1}{r} \right) \right\}_{r=1},$$

and $a_K^{(J-K)}$, $b_K^{(J-K)}$ are derived from functions a_K , b_K which are linear combinations of multipoles having their axes directed along the x and y directions. These latter quantities are determined by the following expansions given by Maxwell:

$$\begin{aligned} a_K &= (K, 0) - \frac{K(K-1)}{1 \cdot 2} (K-2, 2) + \frac{K(K-1)(K-2)(K-3)}{1 \cdot 2 \cdot 3 \cdot 4} (K-4, 4) - \dots, \\ b_K &= K(K-1, 1) - \frac{K(K-1)(K-2)}{1 \cdot 2 \cdot 3} (K-3, 3) \\ &\quad + \frac{K(K-1)(K-2)(K-3)(K-4)}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} (K-5, 5) - \dots \end{aligned}$$

The functions $a_K^{(J-K)}$, $b_K^{(J-K)}$ are derived from a_K and b_K respectively by replacing in the above expansions each factor $(K-L, L)$ by the multipole

$$(K-L, L, J-K) = (-1)^J \frac{\sqrt{(2J+1)}}{J!} \left\{ \frac{d^{K-L}}{dx^{K-L}} \frac{d^L}{dy^L} \frac{d^{J-K}}{dz^{J-K}} \left(\frac{1}{r} \right) \right\}_{r=1}.$$

The functions U_K^J , V_K^J and W^J are all normalized to 4π .

From these formulae we have derived expressions for the real spherical harmonics U_K^J , V_K^J in terms of the multipoles (α, β, γ) (with $\alpha + \beta + \gamma = J$) up to $J = 10$. The results are shown in Table II.

Since the multipoles (α, β, γ) transform like the products $x^\alpha y^\beta z^\gamma$, it is easy to see how these functions transform under the symmetry operations of the tetrahedral group T_d . We denote the linear combinations of the real spherical harmonics U_K^J , V_K^J , W^J which transform according to the irreducible representations A_1 , A_2 , E , F_1 and F_2 of T_d respectively as follows:

$$\left. \begin{aligned} U_\alpha(J) \sim A_1, \quad U_\alpha(J) \sim A_2, \quad \left. \begin{aligned} U_e(J) \\ U_f(J) \end{aligned} \right\} \sim E, \quad \left. \begin{aligned} U_\xi(J) \\ U_\eta(J) \\ U_\zeta(J) \end{aligned} \right\} \sim F_1, \quad \left. \begin{aligned} U_x(J) \\ U_y(J) \\ U_z(J) \end{aligned} \right\} \sim F_2. \end{aligned}$$

If two or more representations of the same type occur in the reduction of D_J , we add superscripts to distinguish the functions, e.g. $U_\alpha^{(1)}(5)$, $U_\alpha^{(2)}(5)$. From the above expressions for the functions in terms of multipoles we found that the combinations given in Table III transform according to irreducible representations as shown. In the case of the threefold degenerate

representations F_1 and F_2 we list only the functions $U_\zeta(J)$, $U_z(J)$ respectively, since these will be sufficient for any perturbation calculation.

As is shown in Table III the functions $U_\zeta(J)$, $U_z(J)$ belonging to the representations F_1 and F_2 are given directly without forming any linear combinations. The general result is that for J even the functions

$$U_z^{(1)}(J), U_\zeta^{(1)}(J), U_z^{(2)}(J), U_\zeta^{(2)}(J), \dots$$

are equal to $V_2^J, V_4^J, V_6^J, V_8^J, \dots$,

respectively, and for J odd the functions

$$U_z^{(1)}(J), U_\zeta^{(1)}(J), U_z^{(2)}(J), U_\zeta^{(2)}(J), \dots$$

are equal respectively to

$$W^J, U_2^J, U_4^J, U_6^J, \dots$$

5. WAVE FUNCTIONS FOR THE CORIOLIS SUBLEVELS

The above reduction enables us at once to form the proper combinations of the vibrational-rotational wave functions of the three Coriolis levels which transform according to irreducible representations of the tetrahedral group. For this purpose we define real wave functions in the same way as we did for the spherical harmonics by

$$U_K(J_L) = \frac{1}{\sqrt{2}} \{W_{-K}(J_L) + (-1)^K W_K(J_L)\},$$

$$V_K(J_L) = \frac{1}{i\sqrt{2}} \{W_{-K}(J_L) - (-1)^K W_K(J_L)\},$$

$$W(J_L) = W_0(J_L),$$

with

$$L = J + 1, J \text{ or } J - 1.$$

For the levels J_{J-1} , J_{J+1} the proper linear combinations of these functions are just the same as for the corresponding spherical harmonics; for the levels J_J , however, which transform according to D_J^g when J is even and according to D_J^g when J is odd, those linear combinations which transform according to F_1 for the spherical harmonics give us the correct linear combinations of the vibrational-rotational wave functions for F_2 and vice versa. Similarly the representations A_1 and A_2 are reversed and also the e - and f -functions of the representation E . In this way we found the linear combinations shown in Table IV. These functions are the correct linear combinations to use in any calculation of a tetrahedrally symmetrical perturbation (except when two or more representations of the same type occur).

TABLE II. REAL SPHERICAL HARMONICS EXPRESSED IN TERMS OF MULTIPOLES

r	K	U'_K	V'_K
1	1	(100)	(010)
2	1	$\frac{2}{\sqrt{3}}(101)$	$\frac{2}{\sqrt{3}}(011)$
2	2	$\frac{1}{\sqrt{3}}\{(200) - (020)\}$	$\frac{2}{\sqrt{3}}(110)$
3	1	$\sqrt{\frac{3}{2}}(102)$	$\sqrt{\frac{3}{2}}(012)$
3	2	$\sqrt{\frac{3}{5}}\{(201) - (021)\}$	$\frac{2\sqrt{3}}{\sqrt{5}}(111)$
3	3	$\frac{1}{\sqrt{10}}\{(300) - 3(120)\}$	$\frac{1}{\sqrt{10}}\{3(210) - (030)\}$
4	1	$\frac{2\sqrt{2}}{\sqrt{5}}(103)$	$\frac{2\sqrt{2}}{\sqrt{5}}(013)$
4	2	$\frac{2}{\sqrt{5}}\{(202) - (022)\}$	$\frac{4}{\sqrt{5}}(112)$
4	3	$\frac{2\sqrt{2}}{\sqrt{35}}\{(301) - 3(121)\}$	$\frac{2\sqrt{2}}{\sqrt{35}}\{3(211) - (031)\}$
4	4	$\frac{1}{\sqrt{35}}\{(400) - 6(220) + (040)\}$	$\frac{4}{\sqrt{35}}\{3(10) - (130)\}$
5	1	$\sqrt{\frac{5}{3}}(104)$	$\sqrt{\frac{5}{3}}(014)$
5	2	$\frac{2\sqrt{5}}{\sqrt{21}}\{(203) - (023)\}$	$\frac{4\sqrt{5}}{\sqrt{21}}(113)$
5	3	$\frac{\sqrt{5}}{\sqrt{14}}\{(302) - 3(122)\}$	$\frac{\sqrt{5}}{\sqrt{14}}\{3(212) - (032)\}$
5	4	$\frac{\sqrt{5}}{3\sqrt{7}}\{(401) - 6(221) + (041)\}$	$\frac{4\sqrt{5}}{3\sqrt{7}}\{3(11) - (131)\}$
5	5	$\frac{1}{3\sqrt{14}}\{(500) - 10(320) + 5(140)\}$	$\frac{1}{3\sqrt{14}}\{5(410) - 10(230) + (050)\}$

6	1	$\frac{2\sqrt{3}}{\sqrt{7}} (105)$	$\frac{2\sqrt{3}}{\sqrt{7}} (015)$
	2	$\frac{\sqrt{15}}{\sqrt{14}} \{(204) - (024)\}$	$\frac{\sqrt{30}}{\sqrt{7}} (114)$
	3	$\frac{\sqrt{10}}{\sqrt{21}} \{(303) - 3 (123)\}$	$\frac{\sqrt{10}}{\sqrt{21}} \{3 (213) - (033)\}$
	4	$\frac{1}{\sqrt{7}} \{(402) - 6 (222) + (042)\}$	$\frac{4}{\sqrt{7}} \{(312) - (132)\}$
	5	$\frac{\sqrt{2}}{\sqrt{77}} \{(501) - 10 (321) + 5 (141)\}$	$-\frac{\sqrt{2}}{\sqrt{77}} \{5 (411) - 10 (231) + (051)\}$
	6	$\frac{1}{\sqrt{(6 \cdot 7 \cdot 11)}} \{(600) - 15 (420) + 15 (240) - (060)\}$	$\frac{\sqrt{2}}{\sqrt{(3 \cdot 7 \cdot 11)}} \{3 (510) - 10 (330) + 3 (150)\}$
7	1	$\frac{\sqrt{7}}{2} (106)$	$\frac{\sqrt{7}}{2} (016)$
	2	$\frac{\sqrt{7}}{\sqrt{6}} \{(205) - (025)\}$	$\frac{\sqrt{14}}{\sqrt{3}} (115)$
	3	$\frac{\sqrt{7}}{2\sqrt{3}} \{(304) - 3 (124)\}$	$\frac{\sqrt{7}}{2\sqrt{3}} \{3 (214) - (034)\}$
	4	$\frac{\sqrt{7}}{\sqrt{33}} \{(403) - 6 (223) + (043)\}$	$\frac{4\sqrt{7}}{\sqrt{33}} \{(313) - (133)\}$
	5	$\frac{\sqrt{7}}{2\sqrt{33}} \{(502) - 10 (322) + 5 (142)\}$	$\frac{\sqrt{7}}{2\sqrt{33}} \{5 (412) - 10 (232) + (052)\}$
	6	$\frac{\sqrt{7}}{\sqrt{(66 \cdot 13)}} \{(601) - 15 (421) + 15 (241) - (061)\}$	$-\frac{\sqrt{14}}{\sqrt{(13 \cdot 33)}} \{3 (511) - 10 (331) + 3 (151)\}$
	7	$\frac{1}{2\sqrt{(13 \cdot 33)}} \{(700) - 21 (520) + 35 (340) - 7 (160)\}$	$\frac{1}{2\sqrt{(13 \cdot 33)}} \{7 (610) - 35 (430) + 21 (250) - (070)\}$

TABLE II (continued)

K	U'_K	V'_K
3		
1	$\frac{4}{3} (107)$	$\frac{4}{3} (017)$
2	$\frac{2\sqrt{14}}{3\sqrt{5}} \{(206) - (026)\}$	$\frac{4\sqrt{14}}{3\sqrt{5}} (116)$
3	$\frac{4\sqrt{7}}{\sqrt{(5.33)}} \{(305) - 3 (125)\}$	$\frac{4\sqrt{7}}{\sqrt{(5.33)}} \{3 (215) - (035)\}$
4	$\frac{2\sqrt{7}}{3\sqrt{11}} \{(404) - 6 (224) + (044)\}$	$\frac{8\sqrt{7}}{3\sqrt{11}} \{(314) - (134)\}$
5	$\frac{4\sqrt{7}}{3\sqrt{(11.13)}} \{(503) - 10 (323) + 5 (143)\}$	$\frac{4\sqrt{7}}{3\sqrt{(11.13)}} \{5 (413) - 10 (233) + (053)\}$
6	$\frac{2\sqrt{2}}{\sqrt{(13.33)}} \{(602) - 15 (422) + 15 (242) - (062)\}$	$\frac{4\sqrt{2}}{\sqrt{(13.33)}} \{3 (512) - 10 (332) + 3 (152)\}$
7	$\frac{4}{3\sqrt{(13.55)}} \{(701) - 21 (521) + 35 (341) - 7 (161)\}$	$\frac{4}{3\sqrt{(13.55)}} \{7 (611) - 35 (431) + 21 (251) - (071)\}$
8	$\frac{1}{3\sqrt{(13.55)}} \{(800) - 28 (620) + 70 (440) - 28 (260) + (080)\}$	$\frac{8}{3\sqrt{(13.55)}} \{(710) - 7 (530) + 7 (350) - (170)\}$
9		
1	$\frac{3}{\sqrt{5}} (108)$	$\frac{3}{\sqrt{5}} (018)$
2	$\frac{6\sqrt{2}}{\sqrt{55}} \{(207) - (027)\}$	$\frac{12\sqrt{2}}{\sqrt{55}} (117)$
3	$\frac{\sqrt{(6.7)}}{\sqrt{55}} \{(306) - 3 (126)\}$	$\frac{\sqrt{(6.7)}}{\sqrt{55}} \{3 (216) - (036)\}$
4	$\frac{6\sqrt{7}}{\sqrt{(13.55)}} \{(405) - 6 (225) + (045)\}$	$\frac{24\sqrt{7}}{\sqrt{(13.55)}} \{(315) - (135)\}$

9	5	$\frac{3\sqrt{2}}{\sqrt{(11.13)}} \{504\} - 10 \{324\} + 5 \{144\}$	$\frac{3\sqrt{2}}{\sqrt{(11.13)}} \{5 \{414\} - 10 \{234\} + \{054\}\}$
	6	$\frac{2\sqrt{6}}{\sqrt{(13.55)}} \{603\} - 15 \{423\} + 15 \{243\} - \{063\}$	$\frac{4\sqrt{6}}{\sqrt{(13.55)}} \{3 \{513\} - 10 \{333\} + 3 \{153\}\}$
	7	$\frac{3}{\sqrt{(13.110)}} \{702\} - 21 \{522\} + 35 \{342\} - 7 \{162\}$	$\frac{3}{\sqrt{(13.110)}} \{7 \{612\} - 35 \{432\} + 21 \{252\} - \{072\}\}$
	8	$\frac{3}{\sqrt{(13.17.55)}} \{801\} - 28 \{621\} + 70 \{441\} - 28 \{261\} + \{081\}$	$\frac{24}{\sqrt{(13.17.55)}} \{711\} - 7 \{531\} + 7 \{351\} - \{171\}$
	9	$\frac{1}{\sqrt{(13.17.110)}} \{900\} - 36 \{720\} + 126 \{540\} - 84 \{360\} + 9 \{180\}$	$\frac{1}{\sqrt{(13.17.110)}} \{9 \{810\} - 84 \{630\} + 126 \{450\} - 36 \{270\} + \{090\}\}$
10	1	$\frac{2\sqrt{5}}{\sqrt{11}} \{109\}$	$\frac{2\sqrt{5}}{\sqrt{11}} \{019\}$
	2	$\frac{\sqrt{15}}{\sqrt{11}} \{208\} - \{028\}$	$\frac{2\sqrt{15}}{\sqrt{11}} \{118\}$
	3	$\frac{2\sqrt{30}}{\sqrt{(11.13)}} \{307\} - 3 \{127\}$	$\frac{2\sqrt{30}}{\sqrt{(11.13)}} \{3 \{217\} - \{037\}\}$
	4	$\frac{2\sqrt{15}}{\sqrt{(11.13)}} \{406\} - 6 \{226\} + \{046\}$	$\frac{8\sqrt{15}}{\sqrt{(11.13)}} \{316\} - \{136\}$
	5	$\frac{2\sqrt{6}}{\sqrt{(11.13)}} \{505\} - 10 \{325\} + 5 \{145\}$	$\frac{2\sqrt{6}}{\sqrt{(11.13)}} \{5 \{415\} - 10 \{235\} + \{055\}\}$
	6	$\frac{\sqrt{15}}{\sqrt{(13.22)}} \{604\} - 15 \{424\} + 15 \{244\} - \{064\}$	$\frac{\sqrt{30}}{\sqrt{(11.13)}} \{3 \{514\} - 10 \{334\} + 3 \{154\}\}$
	7	$\frac{\sqrt{30}}{\sqrt{(11.13.17)}} \{703\} - 21 \{523\} + 35 \{343\} - 7 \{163\}$	$\frac{\sqrt{30}}{\sqrt{(11.13.17)}} \{7 \{613\} - 35 \{433\} + 21 \{253\} - \{073\}\}$
	8	$\frac{\sqrt{5}}{\sqrt{(11.13.17)}} \{802\} - 28 \{622\} + 70 \{442\} - 28 \{262\} + \{082\}$	$\frac{8\sqrt{5}}{\sqrt{(11.13.17)}} \{712\} - 7 \{532\} + 7 \{352\} - \{172\}$
	9	$\frac{\sqrt{10}}{\sqrt{(11.13.17.19)}} \{901\} - 36 \{721\} + 126 \{541\} - 84 \{361\} + 9 \{181\}$	$\frac{\sqrt{10}}{\sqrt{(11.13.17.19)}} \{9 \{811\} - 84 \{631\} + 126 \{451\} - 36 \{271\} + \{091\}\}$
	10	$\frac{1}{\sqrt{(13.17.19.22)}} \{10, 0, 0\} - 45 \{820\} + 210 \{640\} - 210 \{460\} + 45 \{280\} - \{0, 10, 0\}$	$\frac{1}{\sqrt{(13.17.19.22)}} \{10 \{910\} - 120 \{730\} + 252 \{550\} - 120 \{370\} + 10 \{190\}\}$

TABLE III. EXPLICIT REDUCTION OF D_J^g (J EVEN) AND D_J^g (J ODD)

		Linear combination	Multipole
D_1^g	$U_z(1)$	W^1	(001)
D_2^g	$\left\{ \begin{array}{l} U_z(2) \\ U_r(2) \end{array} \right\}$	U_2^2 $- W^2$	$\frac{1}{\sqrt{3}} \{(200) - (020)\}$ $\frac{2}{3} \{(200) + (020) - 2(002)\}$
	$U_z(2)$	V_2^2	$\frac{2}{\sqrt{3}}(110)$
D_3^g	$U_a(3)$	V_2^3	$2\sqrt{\frac{3}{5}}(111)$
	$U_z(3)$	W^3	(003)
	$U_\zeta(3)$	U_2^3	$\sqrt{\frac{3}{5}} \{ \{(201) - (021) \} \}$
D_4^g	$U_a(4)$	$\frac{\sqrt{5}}{2\sqrt{3}} U_4^4 + \frac{\sqrt{7}}{2\sqrt{3}} W^4$	$\frac{2}{\sqrt{21}} \{ (400) + (040) + (004) \}$
	$\left\{ \begin{array}{l} U_a(4) \\ U_r(4) \end{array} \right\}$	$- U_2^4$ $\frac{\sqrt{7}}{2\sqrt{3}} U_4^4 - \frac{\sqrt{5}}{2\sqrt{3}} W^4$	$\frac{2}{\sqrt{5}} \{ (400) - (040) \}$ $\frac{2}{\sqrt{15}} \{ (400) + (040) - 2(004) \}$
	$U_z(4)$	V_2^4	$\frac{4}{\sqrt{5}}(112)$
	$U_\zeta(4)$	V_4^4	$\frac{4}{\sqrt{35}} \{ (310) - (130) \}$

D_6^g	$\left\{ \begin{array}{l} U_e(5) \\ U_f(5) \end{array} \right\}$ $U_z^{(1)}(5)$ $U_t(5)$ $U_z^{(2)}(5)$	V_4^g $-V_2^g$ W^g U_2^g U_4^g	$\frac{4\sqrt{5}}{3\sqrt{7}} \{(311) - (131)\}$ $\frac{4\sqrt{5}}{3\sqrt{21}} \{(311) + (131) - 2(113)\}$ (005) $\frac{2\sqrt{5}}{\sqrt{21}} \{(203) - (023)\}$ $\frac{\sqrt{5}}{3\sqrt{7}} \{(401) - 6(221) + (041)\}$
D_6^g	$U_a(6)$ $U_\alpha(6)$ $\left\{ \begin{array}{l} U_e(6) \\ U_f(6) \end{array} \right\}$ $U_z^{(1)}(6)$ $U_t(6)$ $U_z^{(2)}(6)$	$-\frac{\sqrt{7}}{2\sqrt{2}} U_4^g + \frac{1}{2\sqrt{2}} W^g$ $\frac{\sqrt{11}}{4} U_2^g - \frac{\sqrt{5}}{4} U_6^g$ $\frac{\sqrt{5}}{4} U_2^g + \frac{\sqrt{11}}{4} U_6^g$ $-\frac{1}{2\sqrt{2}} U_4^g - \frac{\sqrt{7}}{2\sqrt{2}} W^g$ V_2^g V_4^g V_6^g	$2\sqrt{2}(222)$ $\frac{2\sqrt{10}}{\sqrt{(3.77)}} \{(420) + (204) + (042) - (402) - (024) - (240)\}$ $\frac{2\sqrt{2}}{\sqrt{21}} \{(600) - (060)\}$ $\frac{2\sqrt{2}}{3\sqrt{7}} \{(600) + (060) - 2(006)\}$ $\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \text{See Table II}$

TABLE III (continued)

		Linear combination	Multipole
D_7^u	$U_a(7)$	$\frac{\sqrt{13}}{2\sqrt{6}} V_2^7 + \frac{\sqrt{11}}{2\sqrt{6}} V_6^7$	$\frac{4\sqrt{7}}{3\sqrt{13}} \{(511) + (151) + (115)\}$
	$\left\{ \begin{array}{l} U_a(7) \\ U_f(7) \\ U_2^{(3)}(7) \\ U_4^{(3)}(7) \\ U_6^{(3)}(7) \end{array} \right\}$	$-V_4^7$	$\frac{4\sqrt{7}}{\sqrt{33}} \{(511) - (151)\}$
		$-\frac{\sqrt{11}}{2\sqrt{6}} V_2^7 + \frac{\sqrt{13}}{2\sqrt{6}} V_6^7$	$\frac{4\sqrt{7}}{3\sqrt{11}} \{(511) + (151) - 2(115)\}$
		W^7	(007)
		U_2^7	} See Table II
		U_4^7	
		U_6^7	
D_8^u	$U_a(8)$	$\frac{\sqrt{7}}{4\sqrt{3}} U_4^8 + \frac{\sqrt{(5.13)}}{8\sqrt{3}} U_8^8 + \frac{\sqrt{33}}{8} W^8$	$\frac{8}{3\sqrt{33}} \{(800) + (080) + (008)\}$
	$\left\{ \begin{array}{l} U_e^{(3)}(8) \\ U_f^{(3)}(8) \end{array} \right\}$	$-\frac{\sqrt{(7.15)}}{2\sqrt{62}} U_2^8 - \frac{\sqrt{(11.13)}}{2\sqrt{62}} U_6^8$	$\frac{8}{\sqrt{(3.31)}} \{(800) - (080)\}$
		$\frac{\sqrt{77}}{4\sqrt{31}} U_4^8 + \frac{\sqrt{(13.55)}}{8\sqrt{31}} U_8^8 - \frac{\sqrt{31}}{8} W^8$	$\frac{8}{3\sqrt{31}} \{(800) + (080) - 2(008)\}$
	$\left\{ \begin{array}{l} U_e^{(3)}(8) \\ U_f^{(3)}(8) \end{array} \right\}$	$-\frac{\sqrt{(11.13)}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{(7.15)}}{2\sqrt{62}} U_6^8$	$\frac{16\sqrt{7}}{3\sqrt{(13.31.55)}} [8 \{(800) - (080)\} - 31 \{(422) - (242)\}]$
		$-\frac{\sqrt{(5.13)}}{\sqrt{(3.31)}} U_4^8 + \frac{2\sqrt{7}}{\sqrt{(3.31)}} U_8^8$	$\frac{16\sqrt{7}}{3\sqrt{(11.13.15.31)}} [8 \{(800) + (080) - 2(008)\} - 31 \{(422) + (242) - 2(224)\}]$
	$U_2^{(3)}(8)$	V_2^8	} See Table II
	$U_4^{(3)}(8)$	V_4^8	
	$U_6^{(3)}(8)$	V_6^8	
	$U_8^{(3)}(8)$	V_8^8	

$D_0^{\#}$	$U_a(9)$ $U_a(9)$ $\left\{ \begin{array}{l} U_s(9) \\ U_l(9) \end{array} \right\}$ $U_x^{(1)}(9)$ $U_y^{(1)}(9)$ $U_z^{(2)}(9)$ $U_x^{(2)}(9)$ $U_z^{(3)}(9)$	$\frac{\sqrt{3}}{4} V_2^9 - \frac{\sqrt{13}}{4} V_6^9$ $\frac{\sqrt{17}}{2\sqrt{6}} V_4^9 - \frac{\sqrt{7}}{2\sqrt{6}} V_8^9$ $\frac{\sqrt{7}}{2\sqrt{6}} V_4^9 + \frac{\sqrt{17}}{2\sqrt{6}} V_8^9$ $-\frac{\sqrt{13}}{4} V_2^9 - \frac{\sqrt{3}}{4} V_6^9$ W^9 U_2^9 U_4^9 U_6^9 U_8^9	$\frac{16\sqrt{2}}{\sqrt{(11.15)}} \{(711) + (171) + (117)\}$ $\frac{16\sqrt{42}}{\sqrt{(13.17.55)}} \{(531) + (315) + (153) - (513) - (135) - (351)\}$ $\frac{16\sqrt{6}}{\sqrt{(13.55)}} \{(711) - (171)\}$ $\frac{16\sqrt{2}}{\sqrt{(13.55)}} \{(711) + (171) - 2(117)\}$ (009) $\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \text{See Table II}$
$D_{10}^{\#}$	$U_a(10)$ $U_a(10)$ $U_o^{(1)}(10)$ $U_l^{(1)}(10)$ $U_o^{(2)}(10)$ $U_l^{(2)}(10)$ $U_o^{(3)}(10)$ $U_z^{(3)}(10)$	$-\frac{\sqrt{11}}{4\sqrt{2}} U_4^{10} - \frac{\sqrt{(11.17)}}{8\sqrt{6}} U_8^{10} + \frac{\sqrt{(5.13)}}{8\sqrt{6}} W^{10}$ $\frac{\sqrt{(13.19)}}{8\sqrt{6}} U_2^{10} + \frac{\sqrt{19}}{16\sqrt{3}} U_6^{10} - \frac{\sqrt{(5.17)}}{16} U_{10}^{10}$ $\frac{7\sqrt{15}}{8\sqrt{(2.29)}} U_2^{10} + \frac{3\sqrt{(13.15)}}{16\sqrt{29}} U_6^{10} + \frac{\sqrt{(13.17.19)}}{16\sqrt{29}} U_{10}^{10}$ $-\frac{\sqrt{(5.13)}}{4\sqrt{(2.29)}} U_4^{10} - \frac{\sqrt{(5.13.17)}}{8\sqrt{(6.29)}} U_8^{10} - \frac{\sqrt{(11.29)}}{8\sqrt{6}} W^{10}$ $\frac{\sqrt{(13.17)}}{4\sqrt{(3.29)}} U_2^{10} - \frac{11\sqrt{17}}{4\sqrt{(6.29)}} U_6^{10} + \frac{\sqrt{(5.19)}}{4\sqrt{(2.29)}} U_{10}^{10}$ $-\sqrt{\frac{17}{29}} U_4^{10} + \frac{2\sqrt{3}}{\sqrt{29}} U_8^{10}$ V_2^{10} $U_4^{(1)}(10)$ $U_6^{(1)}(10)$ $U_8^{(2)}(10)$ $U_6^{(2)}(10)$ $U_z^{(3)}(10)$	$\frac{8\sqrt{2}}{\sqrt{(13.15)}} \{(10, 0, 0) + (0, 10, 0) + (0, 0, 10)\}$ $\frac{8\sqrt{10}}{\sqrt{(11.13.19)}} \{(820) + (208) + (082) - (802) - (028) - (280)\}$ $\frac{8\sqrt{2}}{\sqrt{(11.29)}} \{(10, 0, 0) - (0, 10, 0)\}$ $\frac{16}{\sqrt{(29.66)}} \{(10, 0, 0) + (0, 10, 0) - 2(0, 0, 10)\}$ $\left. \begin{array}{l} \\ \\ \end{array} \right\} \text{Not evaluated}$ $\left. \begin{array}{l} \\ \\ \end{array} \right\} \text{See Table II}$

TABLE IV. WAVE FUNCTIONS FOR THE CORIOLIS SUBLEVELS

Level	Wave function	Linear combination	Level	Wave function	Linear combination
$0_1 F_2$	$(0_1)_z$	$W(0_1)$	$3_4 A_1$	$(3_4)_a$	$\frac{\sqrt{5}}{2\sqrt{3}} U_4(3_4) + \frac{\sqrt{7}}{2\sqrt{3}} W(3_4)$
$1_0 A_1$	$(1_0)_a$	$W(1_0)$	$3_4 E$	$(3_4)_e$	$-U_2(3_4)$
$1_1 F_1$	$(1_1)_\zeta$	$W(1_1)$	$3_4 F_2$	$(3_4)_z$	$V_2(3_4)$
$1_2 E$	$(1_2)_e$	$U_2(1_2)$	$3_4 F_1$	$(3_4)_\zeta$	$V_4(3_4)$
$1_2 F_2$	$(1_2)_z$	$V_2(1_2)$	$4_3 A_1$	$(4_3)_a$	$V_2(4_3)$
$2_1 F_2$	$(2_1)_z$	$W(2_1)$	$4_3 F_1$	$(4_3)_\zeta$	$U_2(4_3)$
$2_2 E$	$(2_2)_e$	$-W(2_2)$	$4_3 F_2$	$(4_3)_z$	$W(4_3)$
$2_2 F_1$	$(2_2)_\zeta$	$V_2(2_2)$	$4_4 A_2$	$(4_4)_a$	$\frac{\sqrt{5}}{2\sqrt{3}} U_4(4_4) + \frac{\sqrt{7}}{2\sqrt{3}} W(4_4)$
$2_3 A_1$	$(2_3)_a$	$V_2(2_3)$	$4_4 E$	$(4_4)_e$	$\frac{\sqrt{7}}{2\sqrt{3}} U_4(4_4) - \frac{\sqrt{5}}{2\sqrt{3}} W(4_4)$
$2_3 F_1$	$(2_3)_\zeta$	$U_2(2_3)$	$4_4 F_1$	$(4_4)_\zeta$	$V_2(4_4)$
$2_3 F_2$	$(2_3)_z$	$W(2_3)$	$4_4 F_2$	$(4_4)_z$	$V_4(4_4)$
$3_2 E$	$(3_2)_e$	$U_2(3_2)$	$4_5 E$	$(4_5)_e$	$V_4(4_5)$
$3_2 F_2$	$(3_2)_z$	$V_2(3_2)$	$4_5 F_2^{(1)}$	$(4_5)_z^{(1)}$	$W(4_5)$
$3_3 A_2$	$(3_3)_a$	$V_2(3_3)$	$4_5 F_1$	$(4_5)_\zeta$	$U_2(4_5)$
$3_3 F_2$	$(3_3)_z$	$U_2(3_3)$	$4_5 F_2^{(2)}$	$(4_5)_z^{(2)}$	$U_4(4_5)$
$3_3 F_1$	$(3_3)_\zeta$	$W(3_3)$			

$5_4 A_1$	$(5_4)_a$	$\frac{\sqrt{5}}{2\sqrt{3}} U_4(5_4) + \frac{\sqrt{7}}{2\sqrt{3}} W(5_4)$	$6_6 F_1^{(1)}$	$(6_6)_\xi^{(1)}$	$V_2(6_6)$
$5_4 E$	$(5_4)_e$	$-U_2(5_4)$	$6_6 F_2$	$(6_6)_z$	$V_4(6_6)$
$5_4 F_2$	$(5_4)_z$	$V_2(5_4)$	$6_6 F_1^{(2)}$	$(6_6)_\xi^{(2)}$	$V_6(6_6)$
$5_4 F_1$	$(5_4)_\xi$	$V_4(5_4)$	$6_7 A_1$	$(6_7)_a$	$\frac{\sqrt{13}}{2\sqrt{6}} V_2(6_7) + \frac{\sqrt{11}}{2\sqrt{6}} V_6(6_7)$
$5_5 E$	$(5_5)_e$	$-V_2(5_5)$	$6_7 E$	$(6_7)_e$	$-V_4(6_7)$
$5_5 F_1^{(1)}$	$(5_5)_\xi^{(1)}$	$W(5_5)$	$6_7 F_1^{(1)}$	$(6_7)_\xi^{(1)}$	$U_2(6_7)$
$5_5 F_2$	$(5_5)_z$	$U_2(5_5)$	$6_7 F_1^{(2)}$	$(6_7)_\xi^{(2)}$	$U_6(6_7)$
$5_5 F_1^{(2)}$	$(5_5)_\xi^{(2)}$	$U_4(5_5)$	$6_7 F_2^{(1)}$	$(6_7)_z^{(1)}$	$W(6_7)$
$5_6 A_1$	$(5_6)_a$	$-\frac{\sqrt{7}}{2\sqrt{2}} U_4(5_6) + \frac{1}{2\sqrt{2}} W(5_6)$	$6_7 F_2^{(2)}$	$(6_7)_z^{(2)}$	$U_4(6_7)$
$5_6 A_2$	$(5_6)_\alpha$	$\frac{\sqrt{11}}{4} U_2(5_6) - \frac{\sqrt{5}}{4} U_6(5_6)$	$7_6 A_1$	$(7_6)_a$	$-\frac{\sqrt{7}}{2\sqrt{2}} U_4(7_6) + \frac{1}{2\sqrt{2}} W(7_6)$
$5_6 E$	$(5_6)_e$	$\frac{\sqrt{5}}{4} U_2(5_6) + \frac{\sqrt{11}}{4} U_6(5_6)$	$7_6 A_2$	$(7_6)_\alpha$	$\frac{\sqrt{11}}{4} U_2(7_6) - \frac{\sqrt{5}}{4} U_6(7_6)$
$5_6 F_2^{(1)}$	$(5_6)_z^{(1)}$	$V_2(5_6)$	$7_6 E$	$(7_6)_e$	$\frac{\sqrt{5}}{4} U_2(7_6) + \frac{\sqrt{11}}{4} U_6(7_6)$
$5_6 F_1$	$(5_6)_\xi$	$V_4(5_6)$	$7_6 F_1$	$(7_6)_\xi$	$V_4(7_6)$
$5_6 F_2^{(2)}$	$(5_6)_z^{(2)}$	$V_6(5_6)$	$7_6 F_2^{(1)}$	$(7_6)_z^{(1)}$	$V_2(7_6)$
			$7_6 F_2^{(2)}$	$(7_6)_z^{(2)}$	$V_6(7_6)$
$6_5 E$	$(6_5)_e$	$V_4(6_5)$	$7_7 A_2$	$(7_7)_\alpha$	$\frac{\sqrt{13}}{2\sqrt{6}} V_2(7_7) + \frac{\sqrt{11}}{2\sqrt{6}} V_6(7_7)$
$6_5 F_2^{(1)}$	$(6_5)_z^{(1)}$	$W(6_5)$	$7_7 E$	$(7_7)_e$	$-\frac{\sqrt{11}}{2\sqrt{6}} V_2(7_7) + \frac{\sqrt{13}}{2\sqrt{6}} V_6(7_7)$
$6_5 F_1$	$(6_5)_\xi$	$U_2(6_5)$	$7_7 F_1^{(1)}$	$(7_7)_\xi^{(1)}$	$W(7_7)$
$6_5 F_2^{(2)}$	$(6_5)_z^{(2)}$	$U_4(6_5)$	$7_7 F_1^{(2)}$	$(7_7)_\xi^{(2)}$	$U_4(7_7)$
				$(7_7)_z^{(1)}$	$U_2(7_7)$
$6_6 A_2$	$(6_6)_\alpha$	$-\frac{\sqrt{7}}{2\sqrt{2}} U_4(6_6) + \frac{1}{2\sqrt{2}} W(6_6)$		$(7_7)_z^{(2)}$	$U_6(7_7)$
$6_6 A_1$	$(6_6)_a$	$\frac{\sqrt{11}}{4} U_2(6_6) - \frac{\sqrt{5}}{4} U_6(6_6)$			
$6_6 E$	$(6_6)_e$	$-\frac{1}{2\sqrt{2}} U_4(6_6) - \frac{\sqrt{7}}{2\sqrt{2}} W(6_6)$			

TABLE IV (continued)

Level	Wave function	Linear combination	Level	Wave function	Linear combination
$7_8 A_1$	$(7_8)_a$	$\frac{\sqrt{7}}{4\sqrt{3}} U_4(7_8) + \frac{\sqrt{(5.13)}}{8\sqrt{3}} U_8(7_8) + \frac{\sqrt{33}}{8} W(7_8)$	$8_6 F_2^{(1)}$	$(8_6)_a^{(1)}$	$V_4(8_6)$
$7_8 E^{(1)}$	$(7_8)_e^{(1)}$	$-\frac{\sqrt{(7.15)}}{2\sqrt{62}} U_2(7_8) - \frac{\sqrt{(11.13)}}{2\sqrt{62}} U_6(7_8)$	$8_6 F_1^{(2)}$	$(8_6)_e^{(2)}$	$V_6(8_6)$
$7_8 E^{(2)}$	$(7_8)_e^{(2)}$	$-\frac{\sqrt{(11.13)}}{2\sqrt{62}} U_2(7_8) + \frac{\sqrt{(7.15)}}{2\sqrt{62}} U_6(7_8)$	$8_6 F_2^{(2)}$	$(8_6)_e^{(2)}$	$V_8(8_6)$
$7_8 F_2^{(1)}$	$(7_8)_z^{(1)}$	$V_2(7_8)$	$8_9 A_1$	$(8_9)_a$	$\frac{\sqrt{3}}{4} V_2(8_9) - \frac{\sqrt{13}}{4} V_6(8_9)$
$7_8 F_1^{(1)}$	$(7_8)_z^{(1)}$	$V_4(7_8)$	$8_9 A_2$	$(8_9)_a$	$\frac{\sqrt{17}}{2\sqrt{6}} V_4(8_9) - \frac{\sqrt{7}}{2\sqrt{6}} V_8(8_9)$
$7_8 F_2^{(2)}$	$(7_8)_z^{(2)}$	$V_6(7_8)$	$8_9 E$	$(8_9)_e$	$\frac{\sqrt{7}}{2\sqrt{6}} V_4(8_9) + \frac{\sqrt{17}}{2\sqrt{6}} V_8(8_9)$
$7_8 F_1^{(2)}$	$(7_8)_z^{(2)}$	$V_8(7_8)$	$8_9 F_2^{(1)}$	$(8_9)_z^{(1)}$	$W(8_9)$
			$8_9 F_2^{(2)}$	$(8_9)_z^{(2)}$	$U_4(8_9)$
			$8_9 F_3^{(3)}$	$(8_9)_z^{(3)}$	$U_8(8_9)$
			$8_9 F_1^{(1)}$	$(8_9)_z^{(1)}$	$U_2(8_9)$
			$8_9 F_2^{(2)}$	$(8_9)_z^{(2)}$	$U_6(8_9)$
$8_7 A_1$	$(8_7)_a$	$\frac{\sqrt{13}}{2\sqrt{6}} V_2(8_7) + \frac{\sqrt{11}}{2\sqrt{6}} V_6(8_7)$	$9_8 A_1$	$(9_8)_a$	$\frac{\sqrt{7}}{4\sqrt{3}} U_4(9_8) + \frac{\sqrt{(5.13)}}{8\sqrt{3}} U_8(9_8) + \frac{\sqrt{33}}{8} W(9_8)$
$8_7 E$	$(8_7)_e$	$-V_4(8_7)$	$9_8 E^{(1)}$	$(9_8)_e^{(1)}$	$-\frac{\sqrt{(7.15)}}{2\sqrt{62}} U_2(9_8) - \frac{\sqrt{(11.13)}}{2\sqrt{62}} U_6(9_8)$
$8_7 F_1^{(1)}$	$(8_7)_z^{(1)}$	$U_2(8_7)$	$9_8 E^{(2)}$	$(9_8)_e^{(2)}$	$-\frac{\sqrt{(11.13)}}{2\sqrt{62}} U_2(9_8) + \frac{\sqrt{(7.15)}}{2\sqrt{62}} U_6(9_8)$
$8_7 F_2^{(2)}$	$(8_7)_z^{(2)}$	$U_6(8_7)$	$9_8 F_1^{(1)}$	$(9_8)_z^{(1)}$	$V_4(9_8)$
$8_7 F_3^{(3)}$	$(8_7)_z^{(3)}$	$W(8_7)$	$9_8 F_2^{(2)}$	$(9_8)_z^{(2)}$	$V_8(9_8)$
$8_7 F_2^{(2)}$	$(8_7)_z^{(2)}$	$U_4(8_7)$	$9_8 F_1^{(1)}$	$(9_8)_z^{(1)}$	$V_2(9_8)$
$8_8 A_2$	$(8_8)_a$	$\frac{\sqrt{7}}{4\sqrt{3}} U_4(8_8) + \frac{\sqrt{(5.13)}}{8\sqrt{3}} U_8(8_8) + \frac{\sqrt{33}}{8} W(8_8)$	$9_8 F_2^{(2)}$	$(9_8)_z^{(2)}$	$V_6(9_8)$
$8_8 E^{(1)}$	$(8_8)_e^{(1)}$	$\frac{\sqrt{77}}{4\sqrt{31}} U_4(8_8) + \frac{\sqrt{(55.13)}}{8\sqrt{31}} U_8(8_8) - \frac{\sqrt{31}}{8} W(8_8)$			
$8_8 E^{(2)}$	$(8_8)_e^{(2)}$	$-\frac{\sqrt{(5.13)}}{\sqrt{(3.31)}} U_4(8_8) + \frac{2\sqrt{7}}{\sqrt{(3.31)}} U_8(8_8)$			
$8_8 F_1^{(1)}$	$(8_8)_z^{(1)}$	$V_2(8_8)$			

$9_9 A_1$	$(9_9)_a$	$\frac{\sqrt{17}}{2\sqrt{6}} V_4(9_9) - \frac{\sqrt{7}}{2\sqrt{6}} V_8(9_9)$	$10_9 A_1$	$(10_9)_a$	$\frac{\sqrt{3}}{4} V_2(10_9) - \frac{\sqrt{13}}{4} V_6(10_9)$
$9_9 A_2$	$(9_9)_\alpha$	$\frac{\sqrt{3}}{4} V_2(9_9) - \frac{\sqrt{13}}{4} V_6(9_9)$	$10_9 A_2$	$(10_9)_\alpha$	$\frac{\sqrt{17}}{2\sqrt{6}} V_4(10_9) - \frac{\sqrt{7}}{2\sqrt{6}} V_8(10_9)$
$9_9 E$	$(9_9)_e$	$-\frac{\sqrt{13}}{4} V_2(9_9) - \frac{\sqrt{3}}{4} V_6(9_9)$	$10_9 E$	$(10_9)_e$	$\frac{\sqrt{7}}{2\sqrt{6}} V_4(10_9) + \frac{\sqrt{17}}{2\sqrt{6}} V_8(10_9)$
$9_9 F_1^{(1)}$	$(9_9)_\xi^{(1)}$	$W(9_9)$	$10_9 F_1^{(1)}$	$(10_9)_\xi^{(1)}$	$U_2(10_9)$
$9_9 F_1^{(2)}$	$(9_9)_\xi^{(2)}$	$U_4(9_9)$	$10_9 F_1^{(2)}$	$(10_9)_\xi^{(2)}$	$U_6(10_9)$
$9_9 F_1^{(3)}$	$(9_9)_\xi^{(3)}$	$U_8(9_9)$	$10_9 F_2^{(1)}$	$(10_9)_\xi^{(1)}$	$W(10_9)$
$9_9 F_2^{(1)}$	$(9_9)_\xi^{(1)}$	$U_2(9_9)$	$10_9 F_2^{(2)}$	$(10_9)_\xi^{(2)}$	$U_4(10_9)$
$9_9 F_2^{(2)}$	$(9_9)_\xi^{(2)}$	$U_6(9_9)$	$10_9 F_2^{(3)}$	$(10_9)_\xi^{(3)}$	$U_8(10_9)$
$9_{10} A_1$	$(9_{10})_a$	$-\frac{\sqrt{11}}{4\sqrt{2}} U_4(9_{10}) - \frac{\sqrt{77}}{8\sqrt{6}} U_8(9_{10}) + \frac{\sqrt{(5, 13)}}{8\sqrt{6}} W(9_{10})$	$10_{10} A_1$	$(10_{10})_a$	$\frac{\sqrt{(13, 19)}}{8\sqrt{6}} U_2(10_{10}) + \frac{\sqrt{19}}{16\sqrt{3}} U_6(10_{10}) - \frac{\sqrt{(5, 17)}}{16} U_{10}(10_{10})$
$9_{10} A_2$	$(9_{10})_\alpha$	$\frac{\sqrt{(13, 19)}}{8\sqrt{6}} U_2(9_{10}) + \frac{\sqrt{19}}{16\sqrt{3}} U_6(9_{10}) - \frac{\sqrt{(5, 17)}}{16} U_{10}(9_{10})$	$10_{10} A_2$	$(10_{10})_\alpha$	$-\frac{\sqrt{11}}{4\sqrt{2}} U_4(10_{10}) - \frac{\sqrt{(11, 17)}}{8\sqrt{6}} U_8(10_{10}) + \frac{\sqrt{(5, 13)}}{8\sqrt{6}} W(10_{10})$
$9_{10} E^{(1)}$	$(9_{10})_e^{(1)}$	$\frac{7\sqrt{15}}{8\sqrt{(2, 29)}} U_2(9_{10}) + \frac{3\sqrt{(13, 15)}}{16\sqrt{29}} U_6(9_{10}) + \frac{\sqrt{(13, 17, 19)}}{16\sqrt{29}} U_{10}(9_{10})$	$10_{10} E^{(1)}$	$(10_{10})_e^{(1)}$	$-\frac{\sqrt{(5, 13)}}{4\sqrt{(2, 29)}} U_4(10_{10}) - \frac{\sqrt{(5, 13, 17)}}{8\sqrt{(6, 29)}} U_8(10_{10}) - \frac{\sqrt{(11, 29)}}{8\sqrt{6}} W(10_{10})$
$9_{10} E^{(2)}$	$(9_{10})_e^{(2)}$	$\frac{\sqrt{(13, 17)}}{4\sqrt{(3, 29)}} U_2(9_{10}) - \frac{11\sqrt{17}}{4\sqrt{(6, 29)}} U_6(9_{10}) + \frac{\sqrt{(5, 19)}}{4\sqrt{(2, 29)}} U_{10}(9_{10})$	$10_{10} E^{(2)}$	$(10_{10})_e^{(2)}$	$-\frac{\sqrt{17}}{29} U_4(10_{10}) + \frac{2\sqrt{3}}{\sqrt{29}} U_8(10_{10})$
$9_{10} F_1^{(1)}$	$(9_{10})_\xi^{(1)}$	$V_4(9_{10})$	$10_{10} F_1^{(1)}$	$(10_{10})_\xi^{(1)}$	$V_2(10_{10})$
$9_{10} F_1^{(2)}$	$(9_{10})_\xi^{(2)}$	$V_8(9_{10})$	$10_{10} F_1^{(2)}$	$(10_{10})_\xi^{(2)}$	$V_6(10_{10})$
$9_{10} F_2^{(1)}$	$(9_{10})_\xi^{(1)}$	$V_2(9_{10})$	$10_{10} F_2^{(1)}$	$(10_{10})_\xi^{(1)}$	$V_{10}(10_{10})$
$9_{10} F_2^{(2)}$	$(9_{10})_\xi^{(2)}$	$V_6(9_{10})$	$10_{10} F_2^{(2)}$	$(10_{10})_\xi^{(2)}$	$V_4(10_{10})$
$9_{10} F_2^{(3)}$	$(9_{10})_\xi^{(3)}$	$V_{10}(9_{10})$	$10_{10} F_2^{(3)}$	$(10_{10})_\xi^{(3)}$	$V_8(10_{10})$

6. CONCLUSION

The vibrational-rotational wave functions given in Table IV will be used in the second part of this paper to evaluate the energies of the corresponding levels when the Coriolis coupling term $H^{(24)}$ is taken into account. We shall derive general formulae for the matrix elements of $H^{(24)}$ between the real vibrational-rotational functions $U_K(J_L)$, $V_K(J_L)$, $W(J_L)$ of ν_4 and the corresponding real vibrational-rotational wave functions of ν_2 . These will then be used to calculate the matrix elements between the proper linear combinations. When two or more representations of the same type occur, the perturbed levels are found by diagonalization of the corresponding perturbation matrix. It should be pointed out that the functions given in Table IV can be used also in evaluating other perturbations, for example second order Coriolis effects involving anharmonicity, or in calculations of the rotation and vibration of methane in the solid state.

The author is indebted to Sir William Bragg, O.M., P.R.S., and the Managers of the Royal Institution, for the facilities afforded at the Davy-Faraday Laboratory where this research was carried out, and to Dr W. H. J. Childs, without whose valuable advice and discussion the work could not have been undertaken.

SUMMARY

The theoretically possible sublevels into which the rotational levels of a threefold degenerate mode of vibration of methane can be split by a tetrahedrally symmetrical perturbation are given up to the tenth rotational quantum number, together with the correct vibrational-rotational wave functions with which to carry out the perturbation calculation. Good approximations to the true modes of vibration of the methane molecule are given and used to evaluate certain Coriolis terms in the Hamiltonian which couple the twofold and threefold degenerate modes of vibration. Table III which gives the tetrahedrally irreducible harmonic functions and extends calculations of Bethe (1929) from the fourth to the tenth quantum number will be useful in a number of connexions.

REFERENCES

- Bethe, H. A. 1929 *Ann. Phys., Lpz.*, **3**, 133.
— 1933 *Handb. Phys.* **24**, **1**, 551.
Childs, W. H. J. 1936 *Proc. Roy. Soc. A*, **153**, 555.

- Cundy, H. M. 1938 *Proc. Roy. Soc. A*, **164**, 420.
Eckart, C. 1935 *Phys. Rev.* **47**, 552.
Elert, W. 1928 *Z. Phys.* **51**, 8.
Jahn, H. A. 1935 *Ann. Phys., Lpz.*, **23**, 529.
Johnston and Dennison 1935 *Phys. Rev.* **48**, 868.
Maxwell, J. C. 1892 "Electricity and Magnetism", I, chapter IX.
Mecke, R. 1930 *Z. Elektrochem.* **36**, 589.
Nielsen and Nielsen 1935 *Phys. Rev.* **48**, 864.
Teller, E. 1934 *Handb. Jb. chem. Phys.* **9**, II, 43.
Tisza, L. 1933 *Z. Phys.* **82**, 48.
-

A new Coriolis perturbation in the methane spectrum

II. Energy levels

By H. A. JAHN

Davy-Faraday Laboratory, The Royal Institution

(Communicated by Sir William Bragg, O.M., P.R.S.—

Received 3 June 1938)

INTRODUCTION

In the first part of this work (Jahn 1938) we found the correct vibrational-rotational wave functions with which to evaluate any tetrahedrally symmetrical perturbation of the rotational levels of a threefold degenerate mode of vibration of the methane molecule. In this second part we use these wave functions to calculate the perturbation due to the Coriolis term which couples the rotational levels of the one quantum threefold degenerate vibration ν_4 and the rotational levels of the one quantum twofold degenerate vibration ν_2 . This Coriolis term was derived in Part I from the general vibrational-rotational Hamiltonian of the molecule. To carry out the perturbation calculation we shall need those combinations of the vibrational-rotational wave functions of ν_2 which transform according to irreducible representations of the tetrahedral group, and these are derived here. We find also general formulae for the matrix elements in question and use these to calculate explicitly the perturbed rotational energy levels of ν_4 up to the tenth rotational quantum number. The rotational levels of ν_2 are of course correspondingly perturbed, but the calculation of these is reserved for another publication.

1. ROTATIONAL SUBLEVELS OF ν_2

The rotational wave functions of the molecule transform, for even and odd values respectively of the rotational quantum number J , according to the representations D_J^g , D_J^u of the group of all rotations and reflexions. The one quantum vibrational wave functions of ν_2 transform according to the representation E of the tetrahedral group. Thus the sublevels into which the rotational levels of ν_2 can be split by a tetrahedral perturbation are given by the reduction of the product representations

$$E \times D_J^g (J \text{ even}) \quad \text{and} \quad E \times D_J^u (J \text{ odd}).$$

From the reduction of D_J^g and D_J^u into irreducible representations of the tetrahedral group (Table I, Part I) and from the reduction of the products of E with other representations of the tetrahedral group, viz. (Tisza 1933)

$$E \times A_1 = E \times A_2 = E,$$

$$E \times E = A_1 + A_2 + E,$$

$$E \times F_1 = E \times F_2 = F_1 + F_2,$$

it is easy to calculate the number and type of sublevels into which each rotational level of ν_2 can be split. These sublevels are listed in Table I.

TABLE I

J	Rotational sublevels of ν_2
0	E
1	$F_1 + F_2$
2	$A_1 + A_2 + E + F_1 + F_2$
3	$E + 2F_1 + 2F_2$
4	$A_1 + A_2 + 2E + 2F_1 + 2F_2$
5	$A_1 + A_2 + E + 3F_1 + 3F_2$
6	$A_1 + A_2 + 3E + 3F_1 + 3F_2$
7	$A_1 + A_2 + 2E + 4F_1 + 4F_2$
8	$2A_1 + 2A_2 + 3E + 4F_1 + 4F_2$
9	$A_1 + A_2 + 3E + 5F_1 + 5F_2$
10	$2A_1 + 2A_2 + 4E + 5F_1 + 5F_2$

To carry out the reduction explicitly we need first of all the explicit reduction of the product representations of E with the other representations

of the tetrahedral group. We denote quantities which transform according to the irreducible representations of the tetrahedral group as follows

$$\alpha \sim A_1, \quad \alpha \sim A_2, \quad \left. \begin{matrix} e \\ f \end{matrix} \right\} \sim E, \quad \left. \begin{matrix} \xi \\ \eta \\ \zeta \end{matrix} \right\} \sim F_1, \quad \left. \begin{matrix} x \\ y \\ z \end{matrix} \right\} \sim F_2.$$

It is then easy to verify that the explicit reduction of the product representations in question are given by the linear combinations shown in Table II. In the case of the more than one-dimensional representations the linear combinations have been so chosen that they transform *in the same way* as the original quantities, as shown in the last column.

The rotational wave functions u_{MK}^J transform under rotations and reflexions about axes fixed in the molecule, independently of the magnetic quantum number M , like the spherical harmonics Y_K^J . We define real rotational wave functions in the same way as we defined real spherical harmonics (Part I) by

$$U_K^J = \frac{1}{\sqrt{2}} \{w_{-K}^J + (-1)^K w_K^J\},$$

$$V_K^J = \frac{1}{i\sqrt{2}} \{w_{-K}^J - (-1)^K w_K^J\},$$

$$W^J = w_0^J,$$

where again we have omitted the quantum number M . From the tetrahedrally irreducible harmonic functions given in Table III, Part I and the linear combinations given in Table II it is easy to find the proper linear combinations of the vibrational-rotational wave functions of ν_2 which transform according to irreducible representations. These are given up to $J = 10$ in Table III, where the two independent vibrational wave functions of ν_2 are denoted by g_e and g_f . As in Table IV, Part I we list in the cases of the representations E , F_1 and F_2 only one of the degenerate functions, viz. the types (e) , (ξ) and (z) , since these will be sufficient for any perturbation calculation. The wave functions here listed will be useful in calculations of the perturbations of the rotational levels of any vibration of type E (i.e. not merely of those of the fundamental ν_2 but also of those of any overtone sub-vibrational level of type E).

TABLE II

Product	Component	Linear combination	Type
$E \times A_1$	E	$\begin{cases} ea \\ fa \end{cases}$	$\begin{matrix} (e) \\ (f) \end{matrix}$
$E \times A_2$	E	$\begin{cases} f\alpha \\ e\alpha \end{cases}$	$\begin{matrix} (e) \\ (f) \end{matrix}$
$E^{(1)} \times E^{(2)}$	A_1	$\frac{1}{\sqrt{2}}(e^{(1)}e^{(2)} + f^{(1)}f^{(2)})$	(a)
	A_2	$\frac{1}{\sqrt{2}}(e^{(1)}f^{(2)} - f^{(1)}e^{(2)})$	(α)
	E	$\begin{cases} \frac{1}{\sqrt{2}}(e^{(1)}f^{(2)} + f^{(1)}e^{(2)}) \\ \frac{1}{\sqrt{2}}(e^{(1)}e^{(2)} - f^{(1)}f^{(2)}) \end{cases}$	$\begin{matrix} (e) \\ (f) \end{matrix}$
$E \times F_1$	F_1	$\begin{cases} -\frac{\sqrt{3}}{2}e\xi - \frac{1}{2}f\xi \\ \frac{\sqrt{3}}{2}e\eta - \frac{1}{2}f\eta \\ f\xi \end{cases}$	$\begin{matrix} (\xi) \\ (\eta) \\ (\zeta) \end{matrix}$
		$\begin{cases} -\frac{1}{2}e\xi + \frac{\sqrt{3}}{2}f\xi \\ -\frac{1}{2}e\eta - \frac{\sqrt{3}}{2}f\eta \\ e\xi \end{cases}$	$\begin{matrix} (x) \\ (y) \\ (z) \end{matrix}$
	F_2	$\begin{cases} -\frac{1}{2}e\xi + \frac{\sqrt{3}}{2}f\xi \\ -\frac{1}{2}e\eta - \frac{\sqrt{3}}{2}f\eta \\ e\xi \end{cases}$	$\begin{matrix} (x) \\ (y) \\ (z) \end{matrix}$
$E \times F_2$	F_1	$\begin{cases} -\frac{1}{2}ex + \frac{\sqrt{3}}{2}fx \\ -\frac{1}{2}ey - \frac{\sqrt{3}}{2}fy \\ ez \end{cases}$	$\begin{matrix} (\xi) \\ (\eta) \\ (\zeta) \end{matrix}$
		$\begin{cases} -\frac{\sqrt{3}}{2}ex - \frac{1}{2}fx \\ \frac{\sqrt{3}}{2}ey - \frac{1}{2}fy \\ fz \end{cases}$	$\begin{matrix} (x) \\ (y) \\ (z) \end{matrix}$
	F_2	$\begin{cases} -\frac{1}{2}ex + \frac{\sqrt{3}}{2}fx \\ -\frac{1}{2}ey - \frac{\sqrt{3}}{2}fy \\ ez \end{cases}$	$\begin{matrix} (x) \\ (y) \\ (z) \end{matrix}$
		$\begin{cases} -\frac{\sqrt{3}}{2}ex - \frac{1}{2}fx \\ \frac{\sqrt{3}}{2}ey - \frac{1}{2}fy \\ fz \end{cases}$	$\begin{matrix} (x) \\ (y) \\ (z) \end{matrix}$

2. GENERAL FORMULAE FOR THE MATRIX ELEMENTS

We found in Part I the following expression for the Coriolis coupling term $H^{(24)}$:

$$H^{(24)} = -\frac{L_x S_x^{(24)} + L_y S_y^{(24)} + L_z S_z^{(24)}}{I},$$

where

$$S_x^{(24)} = \frac{\hbar}{i} \zeta_{24} \left\{ \frac{1}{2} \left(q_{2e} \frac{\partial}{\partial q_{4x}} - q_{4x} \frac{\partial}{\partial q_{2e}} \right) - \frac{\sqrt{3}}{2} \left(q_{2f} \frac{\partial}{\partial q_{4x}} - q_{4x} \frac{\partial}{\partial q_{2f}} \right) \right\},$$

$$S_y^{(24)} = \frac{\hbar}{i} \zeta_{24} \left\{ \frac{1}{2} \left(q_{2e} \frac{\partial}{\partial q_{4y}} - q_{4y} \frac{\partial}{\partial q_{2e}} \right) + \frac{\sqrt{3}}{2} \left(q_{2f} \frac{\partial}{\partial q_{4y}} - q_{4y} \frac{\partial}{\partial q_{2f}} \right) \right\},$$

$$S_z^{(24)} = \frac{\hbar}{i} \zeta_{24} \left\{ - \left(q_{2e} \frac{\partial}{\partial q_{4z}} - q_{4z} \frac{\partial}{\partial q_{2e}} \right) \right\}.$$

It is convenient to replace the normal co-ordinates by dimensionless quantities defined by

$$x = \sqrt{\frac{mw_4}{\hbar}} q_{4x}, \quad y = \sqrt{\frac{mw_4}{\hbar}} q_{4y}, \quad z = \sqrt{\frac{mw_4}{\hbar}} q_{4z},$$

$$e = \sqrt{\frac{mw_2}{\hbar}} q_{2e}, \quad f = \sqrt{\frac{mw_2}{\hbar}} q_{2f}.$$

Putting $\alpha = \sqrt{\frac{w_4}{w_2}}$, $\beta = \sqrt{\frac{w_2}{w_4}}$, one finds easily that

$$S_x^{(24)} = \frac{\hbar}{i} \zeta_{24} \left\{ \frac{1}{2} \left(\alpha e \frac{\partial}{\partial x} - \beta x \frac{\partial}{\partial e} \right) - \frac{\sqrt{3}}{2} \left(\alpha f \frac{\partial}{\partial x} - \beta x \frac{\partial}{\partial f} \right) \right\},$$

$$S_y^{(24)} = \frac{\hbar}{i} \zeta_{24} \left\{ \frac{1}{2} \left(\alpha e \frac{\partial}{\partial y} - \beta y \frac{\partial}{\partial e} \right) + \frac{\sqrt{3}}{2} \left(\alpha f \frac{\partial}{\partial y} - \beta y \frac{\partial}{\partial f} \right) \right\},$$

$$S_z^{(24)} = \frac{\hbar}{i} \zeta_{24} \left\{ - \left(\alpha e \frac{\partial}{\partial z} - \beta z \frac{\partial}{\partial e} \right) \right\}.$$

Now the normalized wave functions of a harmonic oscillator are given by

$$\psi_n = e^{-\frac{1}{2}x^2} \phi_n \quad \text{with} \quad \phi_n = \frac{H_n(x)}{\sqrt{\{2^n(n!)\}}},$$

where $H_n(x)$ is the n th Hermitean polynome. Here again $x = \sqrt{\frac{mw}{\hbar}} q$, where

w is the frequency, m the reduced mass and q the normal co-ordinate. For these wave functions we have

$$x\psi_n = \sqrt{\frac{n+1}{2}}\psi_{n+1} + \sqrt{\frac{n}{2}}\psi_{n-1},$$

$$\frac{d}{dx}\psi_n = -\sqrt{\frac{n+1}{2}}\psi_{n+1} + \sqrt{\frac{n}{2}}\psi_{n-1}.$$

Let us denote the zero-point wave function ψ_0 of the vibration ν_2 by g_0 and that of the vibration ν_4 by f_0 . The two independent one quantum vibrational wave functions ψ_1 of ν_2 we denote by g_e and g_f , and the three independent one quantum vibrational wave functions of ν_4 by f_x , f_y and f_z . From the formulae just given we find then

$$eg_0 = \frac{1}{\sqrt{2}}g_e, \quad \frac{\partial}{\partial e}g_0 = -\frac{1}{\sqrt{2}}g_e,$$

$$fg_0 = \frac{1}{\sqrt{2}}g_f, \quad \frac{\partial}{\partial f}g_0 = -\frac{1}{\sqrt{2}}g_f,$$

$$xf_x = yf_y = zf_z = \frac{1}{\sqrt{2}}f_0,$$

$$\frac{d}{dx}f_x = \frac{d}{dy}f_y = \frac{d}{dz}f_z = \frac{1}{\sqrt{2}}f_0.$$

From these one calculates easily

$$\left\{\frac{1}{2}\left(\alpha e \frac{\partial}{\partial x} - \beta x \frac{\partial}{\partial e}\right) - \frac{\sqrt{3}}{2}\left(\alpha f \frac{\partial}{\partial x} - \beta x \frac{\partial}{\partial f}\right)\right\}f_x g_0 = \frac{\alpha + \beta}{4}(g_e - \sqrt{3}g_f)f_0,$$

$$\left\{\frac{1}{2}\left(\alpha e \frac{\partial}{\partial y} - \beta y \frac{\partial}{\partial e}\right) + \frac{\sqrt{3}}{2}\left(\alpha f \frac{\partial}{\partial y} - \beta y \frac{\partial}{\partial f}\right)\right\}f_y g_0 = \frac{\alpha + \beta}{4}(g_e + \sqrt{3}g_f)f_0,$$

$$-\left(\alpha e \frac{\partial}{\partial z} - \beta z \frac{\partial}{\partial e}\right)f_z g_0 = -\frac{\alpha + \beta}{2}g_e f_0.$$

Thus putting

$$s_x = \frac{i}{\hbar} \frac{1}{\zeta_{24}} \frac{4}{\alpha + \beta} S_x^{(24)}, \quad s_y = \frac{i}{\hbar} \frac{1}{\zeta_{24}} \frac{4}{\alpha + \beta} S_y^{(24)}, \quad s_z = \frac{i}{\hbar} \frac{1}{\zeta_{24}} \frac{4}{\alpha + \beta} S_z^{(24)},$$

we find, omitting the zero-point wave functions,

$$s_x f_x = g_e - \sqrt{3}g_f,$$

$$s_y f_y = g_e + \sqrt{3}g_f,$$

$$s_z f_z = -2g_e.$$

If we put
$$L_x = \frac{\hbar}{i} l_x, \quad L_y = \frac{\hbar}{i} l_y, \quad L_z = \frac{\hbar}{i} l_z,$$

we have then
$$H^{(24)} = \frac{\hbar^2}{I} \zeta_{24} \frac{\alpha + \beta}{4} A,$$

where
$$A = l_x s_x + l_y s_y + l_z s_z.$$

The squares of the matrix elements of $H^{(24)}$ are then given in terms of the squares of the matrix elements of A by

$$|H^{(24)}|^2 = \left(\frac{\hbar^2}{I}\right)^2 (\zeta_{24})^2 \frac{(\alpha + \beta)^2}{16} |A|^2.$$

Putting in the numerical factors ($\hbar^2/I = 2B = 10.5 \text{ cm.}^{-1}$, $\zeta_{24}^2 = 0.4$, $(\alpha + \beta)^2 = 4.0$), we find

$$|H^{(24)}|^2 = 11.0 |A|^2 (\text{cm.}^{-1})^2.$$

We proceed now to deduce general formulae for the matrix elements of the operator A . Putting

$$\begin{aligned} l_1 &= l_x + i l_y, & s_1 &= s_x + i s_y, \\ l_{-1} &= l_x - i l_y, & s_{-1} &= s_x - i s_y, \\ l_0 &= l_z, & s_0 &= s_z, \end{aligned}$$

we have
$$A = \frac{1}{2}(l_1 s_{-1} + l_{-1} s_1) + l_0 s_0.$$

Putting as before (Part I)

$$v_1 = -\frac{f_x - i f_y}{\sqrt{2}}, \quad v_0 = f_z, \quad v_{-1} = \frac{f_x + i f_y}{\sqrt{2}},$$

we find easily
$$\begin{aligned} s_1 v_1 &= -\sqrt{2} g_e, & s_{-1} v_1 &= \sqrt{6} g_f, \\ s_1 v_{-1} &= -\sqrt{6} g_f, & s_{-1} v_{-1} &= \sqrt{2} g_e, \\ s_0 v_0 &= -2 g_e, \end{aligned}$$

the other products not given being zero. Further we have (cf. Jahn 1935)

$$\begin{aligned} l_1 w_K^J &= i \sqrt{\frac{1}{2}} (J + K) (J - K + 1) w_{K-1}^J, \\ l_{-1} w_K^J &= i \sqrt{\frac{1}{2}} (J - K) (J + K + 1) w_{K+1}^J, \\ l_0 w_K^J &= i K w_K^J. \end{aligned}$$

From the general formulae given in Part I for the vibrational-rotational wave functions $W_K(J_{J-1})$, $W_K(J_J)$ and $W_K(J_{J+1})$ of the three Coriolis levels

of ν_4 we find the following general formulae for the matrix elements of the operator A . In the case of the functions $W_K(J_J)$ a common phase factor i has been omitted.

General formulae for the matrix elements

R branch (J_{J-1}) levels.

$$AW(J_{J-1}) = -\sqrt{3} \sqrt{\frac{(J-1)(J+1)(J+2)}{2(2J+1)}} g_f V_2^J,$$

$$\begin{aligned} AU_K(J_{J-1}) &= 3K \sqrt{\frac{(J+K)(J-K)}{J(2J+1)}} g_e V_K^J \\ &+ \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(2J+1)\}}} \sqrt{\{(J-K)(J-K+1)(J+K-1)(J-K+2)\}} g_f V_{K-2}^J \\ &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(2J+1)\}}} \sqrt{\{(J+K)(J+K+1)(J-K-1)(J+K+2)\}} g_f V_{K+2}^J, \end{aligned}$$

$$\begin{aligned} AV_K(J_{J-1}) &= 3K \sqrt{\frac{(J+K)(J-K)}{J(2J+1)}} g_e U_K^J \\ &+ \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(2J+1)\}}} \sqrt{\{(J-K)(J-K+1)(J+K-1)(J-K+2)\}} g_f U_{K-2}^J \\ &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(2J+1)\}}} \sqrt{\{(J+K)(J+K+1)(J-K-1)(J+K+2)\}} g_f U_{K+2}^J. \end{aligned}$$

Q branch (J_J levels).

$$AW(J_J) = \sqrt{\{J(J+1)\}} g_e W^J - \sqrt{\frac{3}{2}} \sqrt{\{(J-1)(J+2)\}} g_f U_2^J,$$

$$\begin{aligned} AU_K(J_J) &= \frac{J(J+1) - 3K^2}{\sqrt{\{J(J+1)\}}} g_e U_K^J \\ &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(J+1)\}}} \sqrt{\{(J+K)(J-K+1)(J+K-1)(J-K+2)\}} g_f U_{K-2}^J \\ &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(J+1)\}}} \sqrt{\{(J-K)(J+K+1)(J-K-1)(J+K+2)\}} g_f U_{K+2}^J, \end{aligned}$$

$$\begin{aligned} AV_K(J_J) &= \frac{J(J+1) - 3K^2}{\sqrt{\{J(J+1)\}}} g_e V_K^J \\ &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(J+1)\}}} \sqrt{\{(J+K)(J-K+1)(J+K-1)(J-K+2)\}} g_f V_{K-2}^J \\ &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{J(J+1)\}}} \sqrt{\{(J-K)(J+K+1)(J-K-1)(J+K+2)\}} g_f V_{K+2}^J. \end{aligned}$$

P branch (J_{J+1} levels).

$$\begin{aligned}
 AW(J_{J+1}) &= -\sqrt{3} \sqrt{\frac{J(J-1)(J+2)}{2(2J+1)}} g_f V_2^J, \\
 AU_K(J_{J+1}) &= -3K \sqrt{\frac{(J+K+1)(J-K+1)}{(J+1)(2J+1)}} g_e V_K^J \\
 &+ \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{(J+1)(2J+1)\}}} \sqrt{\{(J+K+1)(J+K-1)(J-K+2)(J+K)\}} g_f V_{K-2}^J \\
 &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{(J+1)(2J+1)\}}} \sqrt{\{(J-K-1)(J-K+1)(J+K+2)(J-K)\}} g_f V_{K+2}^J, \\
 AV_K(J_{J+1}) &= -3K \sqrt{\frac{(J+K+1)(J-K+1)}{(J+1)(2J+1)}} g_e U_K^J \\
 &+ \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{(J+1)(2J+1)\}}} \sqrt{\{(J+K+1)(J+K-1)(J-K+2)(J+K)\}} g_f U_{K-2}^J \\
 &- \frac{\sqrt{3}}{2} \frac{1}{\sqrt{\{(J+1)(2J+1)\}}} \sqrt{\{(J-K-1)(J-K+1)(J+K+2)(J-K)\}} g_f U_{K+2}^J,
 \end{aligned}$$

where $V_0^J = 0, \quad U_0^J = \sqrt{2} W^J.$

3. MATRIX ELEMENTS FOR THE SUBLEVELS

From the general formulae given above it is possible to calculate the matrix elements of the operator A between the wave functions for the Coriolis sublevels of ν_4 listed in Table IV, Part I and the wave functions for the corresponding rotational sublevels of ν_2 given in Table III below. These matrix elements have been evaluated up to $J = 10$ and are listed in Table IV below. Those matrix elements which are not listed vanish.

As an example as to how the matrix elements listed in Table IV were derived and to explain the notation, we take the level $7_7 E$. We find from Table IV, Part I

$$(7_7)_e = -\frac{\sqrt{11}}{2\sqrt{6}} V_2(7_7) + \frac{\sqrt{13}}{2\sqrt{6}} V_6(7_7).$$

Hence from the general formulae given above we find

$$\begin{aligned}
 A(7_7)_e &= -\frac{\sqrt{11}}{2\sqrt{6}} AV_2(7_7) + \frac{\sqrt{13}}{2\sqrt{6}} AV_6(7_7) \\
 &= -\frac{11\sqrt{11}}{2\sqrt{21}} g_e V_2^7 - \frac{13\sqrt{13}}{2\sqrt{21}} g_e V_6^7 + \frac{2\sqrt{2}}{\sqrt{7}} g_f V_4^7.
 \end{aligned}$$

This expression must be a linear combination of the functions $(7)_e^{(1)}$, $(7)_e^{(2)}$ which according to Table III are defined by

$$(7)_e^{(1)} = \frac{\sqrt{13}}{2\sqrt{6}} g_e V_2^7 + \frac{\sqrt{11}}{2\sqrt{6}} g_e V_6^7,$$

$$(7)_e^{(2)} = -\frac{\sqrt{11}}{4\sqrt{3}} g_e V_2^7 + \frac{\sqrt{13}}{4\sqrt{3}} g_e V_6^7 - \frac{1}{\sqrt{2}} g_f V_4^7.$$

Putting

$$A(7)_e = a_1(7)_e^{(1)} + a_2(7)_e^{(2)},$$

we find the following equations for a_1 and a_2 :

$$\frac{\sqrt{13}}{2\sqrt{6}} a_1 - \frac{\sqrt{11}}{4\sqrt{3}} a_2 = -\frac{11\sqrt{11}}{2\sqrt{21}},$$

$$\frac{\sqrt{11}}{2\sqrt{6}} a_1 + \frac{\sqrt{13}}{4\sqrt{3}} a_2 = -\frac{13\sqrt{13}}{2\sqrt{21}},$$

$$-\frac{a_2}{\sqrt{2}} = \frac{2\sqrt{2}}{\sqrt{7}}.$$

These equations are consistent and give the values

$$a_1 = -\frac{\sqrt{(13 \cdot 22)}}{\sqrt{7}}, \quad a_2 = -\frac{4}{\sqrt{7}},$$

i.e.

$$A(7)_e = -\frac{\sqrt{(13 \cdot 22)}}{\sqrt{7}} (7)_e^{(1)} - \frac{4}{\sqrt{7}} (7)_e^{(2)}.$$

This equation is expressed in the Table by putting

$$a_1 = \{(7)_e | A | (7)_e^{(1)}\} = -\frac{\sqrt{(13 \cdot 22)}}{\sqrt{7}},$$

$$a_2 = \{(7)_e | A | (7)_e^{(2)}\} = -\frac{4}{\sqrt{7}}.$$

In every case the consistency of the equations was verified as a check on the calculations. A further check will be mentioned below. When two or more sublevels of the same symmetry type occur in a rotational level of ν_4 we use additional suffixes. Thus in the case of the levels $6_5 2F_2$ we put

$$a_{12} = \{(6_5)_z^{(1)} | A | (6)_z^{(2)}\}, \quad a_{21} = \{(6_5)_z^{(2)} | A | (6)_z^{(1)}\},$$

$$a_{22} = \{(6_5)_z^{(2)} | A | (6)_z^{(2)}\}, \quad a_{23} = \{(6_5)_z^{(2)} | A | (6)_z^{(3)}\},$$

the matrix elements a_{11} , a_{13} being zero.

4. ENERGIES OF THE SUBLEVELS OF ν_4

To calculate the energies of the individual sublevels of ν_4 we apply second order perturbation theory. If there is no accidental degeneracy, i.e. if no two sublevels of the same symmetry type occur, the result is well known: denoting by E_4 the unperturbed energy of the rotational level of ν_4 , where E_4 takes into account the splitting caused by Teller's Coriolis term $H^{(4)}$, and by E_2 the unperturbed energy of the rotational level of ν_2 with which the level of ν_4 is coupled, then the perturbed energies of the two levels are given by the roots of the determinant

$$\begin{vmatrix} \lambda - E_4 & -c \\ -c & \lambda - E_2 \end{vmatrix} = 0,$$

where c is the coupling matrix element. Putting $\lambda - E_4 = x$, $\Delta = E_2 - E_4$, we find

$$x(x - \Delta) - c^2 = 0,$$

which gives, neglecting x^2 compared with Δx ,

$$x = -\frac{c^2}{\Delta} = -\frac{c^2}{E_2 - E_4}.$$

If there is a twofold accidental degeneracy in both the perturbing levels, then the two perturbed levels of ν_4 are given by the two lowest roots of the determinantal equation

$$\begin{vmatrix} \lambda - E_4 & 0 & -c_{11} & -c_{12} \\ 0 & \lambda - E_4 & -c_{21} & -c_{22} \\ \hline -c_{11} & -c_{21} & \lambda - E_2 & 0 \\ -c_{12} & -c_{22} & 0 & \lambda - E_2 \end{vmatrix} = 0.$$

Putting again $\lambda - E_4 = x$, $E_2 - E_4 = \Delta$,

we find
$$x^2(x - \Delta)^2 - Rx(x - \Delta) + D^2 = 0,$$

where

$$R = c_{11}^2 + c_{12}^2 + c_{21}^2 + c_{22}^2,$$

$$D = c_{11}c_{22} - c_{12}c_{21}.$$

Neglecting Rx^2 compared with Δ^2x^2 , we find the quadratic equation

$$x^2\Delta^2 + R\Delta x + D^2 = 0,$$

TABLE III. WAVE FUNCTIONS FOR THE ROTATIONAL SUBLEVELS OF ν_2

J	Sub-level	Wave function	Linear combination	J	Sub-level	Wave function	Linear combination
0	E	$(0)_e$	$g_e W^0$	4	$E^{(1)}$	$(4)_e^{(1)}$	$g_e \left(\frac{\sqrt{5}}{2\sqrt{3}} U_4^4 + \frac{\sqrt{7}}{2\sqrt{3}} W^4 \right)$
1	F_1 F_2	$(1)_e$ $(1)_z$	$g_e W^1$ $g_f W^1$		$E^{(2)}$	$(4)_e^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{7}}{2\sqrt{3}} U_4^4 - \frac{\sqrt{5}}{2\sqrt{3}} W^4 \right) - g_f U_2^4 \right\}$
2	A_1 A_2	$(2)_e$ $(2)_z$	$\frac{1}{\sqrt{2}} (g_e U_2^2 - g_f W^2)$ $-\frac{1}{\sqrt{2}} (g_e W^2 + g_f U_2^2)$		$F_1^{(1)}$ $F_1^{(2)}$ $F_2^{(1)}$ $F_2^{(2)}$	$(4)_e^{(1)}$ $(4)_e^{(2)}$ $(4)_z^{(1)}$ $(4)_z^{(2)}$	$g_f V_4^4$ $g_e V_2^4$ $g_e V_4^4$ $g_f V_2^4$
3	E F_1 F_2	$(2)_e$ $(2)_e$ $(2)_z$	$\frac{1}{\sqrt{2}} (-g_e W^2 + g_f U_2^2)$ $g_e V_2^2$ $g_f V_2^2$	5	A_1	$(5)_e$	$\frac{1}{\sqrt{2}} (g_e V_4^5 - g_f V_2^5)$
					A_2	$(5)_e$	$-\frac{1}{\sqrt{2}} (g_e V_2^5 + g_f V_4^5)$
					E	$(5)_e$	$\frac{1}{\sqrt{2}} (-g_e V_2^5 + g_f V_4^5)$
					$F_1^{(1)}$ $F_1^{(2)}$ $F_2^{(1)}$ $F_2^{(2)}$	$(5)_e^{(1)}$ $(5)_e^{(2)}$ $(5)_z^{(1)}$ $(5)_z^{(2)}$	$g_f U_2^5$ $g_e W^5$ $g_e U_4^5$ $g_e U_2^5$
4	A_1 A_2	$(4)_e$ $(4)_z$	$\frac{1}{\sqrt{2}} \left\{ -g_e U_2^4 + g_f \left(\frac{\sqrt{7}}{2\sqrt{3}} U_4^4 - \frac{\sqrt{5}}{2\sqrt{3}} W^4 \right) \right\}$ $\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{7}}{2\sqrt{3}} U_4^4 - \frac{\sqrt{5}}{2\sqrt{3}} W^4 \right) + g_f U_2^4 \right\}$		$F_1^{(3)}$ $F_2^{(3)}$	$(5)_z^{(3)}$ $(5)_z^{(3)}$	$g_f W^5$ $g_f U_4^5$

6	A_1	$(6)_a$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{5}}{4} U_2^6 + \frac{\sqrt{11}}{4} U_6^6 \right) - g_f \left(\frac{1}{2\sqrt{2}} U_4^6 + \frac{\sqrt{7}}{2\sqrt{2}} W_6^6 \right) \right\}$	7	$E^{(2)}$	$(7)_e^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\frac{\sqrt{11}}{2\sqrt{6}} V_2^7 + \frac{\sqrt{13}}{2\sqrt{6}} V_6^7 \right) - g_f V_4^7 \right\}$
	A_2	$(6)_a$	$-\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{1}{2\sqrt{2}} U_4^6 + \frac{\sqrt{7}}{2\sqrt{2}} W_6^6 \right) + g_f \left(\frac{\sqrt{5}}{4} U_2^6 + \frac{\sqrt{11}}{4} U_6^6 \right) \right\}$		$F_1^{(1)}$	$(7)_e^{(1)}$	$g_f U_2^7$
	$E^{(1)}$	$(6)_e^{(1)}$	$g_e \left(-\frac{\sqrt{7}}{2\sqrt{2}} U_4^6 + \frac{1}{2\sqrt{2}} W_6^6 \right)$		$F_1^{(2)}$	$(7)_e^{(2)}$	$g_f U_6^7$
	$E^{(2)}$	$(6)_e^{(2)}$	$g_f \left(\frac{\sqrt{11}}{4} U_2^6 - \frac{\sqrt{5}}{4} U_6^6 \right)$		$F_1^{(3)}$	$(7)_e^{(3)}$	$g_e W^7$
	$E^{(3)}$	$(6)_e^{(3)}$	$\frac{1}{\sqrt{2}} \left\{ -g_e \left(\frac{1}{2\sqrt{2}} U_4^6 + \frac{\sqrt{7}}{2\sqrt{2}} W_6^6 \right) + g_f \left(\frac{\sqrt{5}}{4} U_2^6 + \frac{\sqrt{11}}{4} U_6^6 \right) \right\}$		$F_1^{(4)}$	$(7)_e^{(4)}$	$g_e U_4^7$
	$F_1^{(1)}$	$(6)_e^{(1)}$	$g_f V_4^6$		$F_2^{(1)}$	$(7)_e^{(1)}$	$g_e U_2^7$
	$F_1^{(2)}$	$(6)_e^{(2)}$	$g_e V_2^6$		$F_2^{(2)}$	$(7)_e^{(2)}$	$g_e U_6^7$
	$F_1^{(3)}$	$(6)_e^{(3)}$	$g_e V_6^6$		$F_2^{(3)}$	$(7)_e^{(3)}$	$g_f W^7$
	$F_2^{(1)}$	$(6)_e^{(1)}$	$g_e V_4^6$		$F_2^{(4)}$	$(7)_e^{(4)}$	$g_f U_4^7$
	$F_2^{(2)}$	$(6)_e^{(2)}$	$g_f V_2^6$				
	$F_2^{(3)}$	$(6)_e^{(3)}$	$g_f V_6^6$				
7	A_1	$(7)_a$	$\frac{1}{\sqrt{2}} \left\{ -g_e V_4^7 + g_f \left(-\frac{\sqrt{11}}{2\sqrt{6}} V_2^7 + \frac{\sqrt{13}}{2\sqrt{6}} V_6^7 \right) \right\}$	8	$A_1^{(1)}$	$(8)_a^{(1)}$	$\frac{1}{\sqrt{2}} \left\{ -g_e \left(\frac{\sqrt{7.15}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{11.13}}{2\sqrt{62}} U_6^8 \right) + g_f \left(\frac{\sqrt{77}}{4\sqrt{31}} U_4^8 + \frac{\sqrt{13.55}}{8\sqrt{31}} U_8^8 - \frac{\sqrt{31}}{8} W^8 \right) \right\}$
	A_2	$(7)_a$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\frac{\sqrt{11}}{2\sqrt{6}} V_2^7 + \frac{\sqrt{13}}{2\sqrt{6}} V_6^7 \right) + g_f V_4^7 \right\}$		$A_1^{(2)}$	$(8)_a^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\frac{\sqrt{11.13}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{7.15}}{2\sqrt{62}} U_6^8 \right) + g_f \left(-\frac{\sqrt{5.13}}{\sqrt{3.31}} U_4^8 + \frac{2\sqrt{7}}{\sqrt{3.31}} U_8^8 \right) \right\}$
	$E^{(1)}$	$(7)_e^{(1)}$	$g_e \left(\frac{\sqrt{13}}{2\sqrt{6}} V_2^7 + \frac{\sqrt{11}}{2\sqrt{6}} V_6^7 \right)$		$A_2^{(1)}$	$(8)_a^{(1)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{77}}{4\sqrt{31}} U_4^8 + \frac{\sqrt{13.55}}{8\sqrt{31}} U_8^8 - \frac{\sqrt{31}}{8} W^8 \right) + g_f \left(\frac{\sqrt{7.15}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{11.13}}{2\sqrt{62}} U_6^8 \right) \right\}$

TABLE III (contd.)

J	Sub- level	Wave func- tion	Linear combination	J	Sub- level	Wave func- tion	Linear combination
8	$A_2^{(2)}$	$(8)_{\alpha}^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\sqrt{\frac{5.13}{3.31}} U_4^8 + \frac{2\sqrt{7}}{\sqrt{3.31}} U_8^8 \right) - g_f \left(-\frac{\sqrt{11.13}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{7.15}}{2\sqrt{62}} U_6^8 \right) \right\}$	9	A_1	$(9)_{\alpha}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{7}}{2\sqrt{6}} V_4^8 + \frac{\sqrt{17}}{2\sqrt{6}} V_8^8 \right) - g_f \left(\frac{\sqrt{13}}{4} V_2^8 + \frac{\sqrt{3}}{4} V_6^8 \right) \right\}$
	$E^{(1)}$	$(8)_{\alpha}^{(1)}$	$g_e \left(\frac{\sqrt{7}}{4\sqrt{3}} U_4^8 + \frac{\sqrt{5.13}}{8\sqrt{3}} U_8^8 + \frac{\sqrt{33}}{8} W^8 \right)$		A_2	$(9)_{\alpha}$	$\frac{1}{\sqrt{2}} \left\{ -g_e \left(\frac{\sqrt{13}}{4} V_2^8 + \frac{\sqrt{3}}{4} V_6^8 \right) - g_f \left(\frac{\sqrt{7}}{2\sqrt{6}} V_4^8 + \frac{\sqrt{17}}{2\sqrt{6}} V_8^8 \right) \right\}$
	$E^{(2)}$	$(8)_{\alpha}^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{77}}{4\sqrt{31}} U_4^8 + \frac{\sqrt{13.55}}{8\sqrt{31}} U_8^8 - \frac{\sqrt{31}}{8} W^8 \right) - g_f \left(\frac{\sqrt{7.15}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{11.13}}{2\sqrt{62}} U_6^8 \right) \right\}$		$E^{(1)}$	$(9)_{\alpha}^{(1)}$	$g_e \left(\frac{\sqrt{3}}{4} V_2^8 - \frac{\sqrt{13}}{4} V_6^8 \right)$
	$E^{(3)}$	$(8)_{\alpha}^{(3)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\sqrt{\frac{5.13}{3.31}} U_4^8 + \frac{2\sqrt{7}}{\sqrt{3.31}} U_8^8 \right) + g_f \left(-\frac{\sqrt{11.13}}{2\sqrt{62}} U_2^8 + \frac{\sqrt{7.15}}{2\sqrt{62}} U_6^8 \right) \right\}$		$E^{(2)}$	$(9)_{\alpha}^{(2)}$	$g_f \left(\frac{\sqrt{17}}{2\sqrt{6}} V_4^8 - \frac{\sqrt{7}}{2\sqrt{6}} V_8^8 \right)$
	$F_1^{(1)}$	$(8)_{\alpha}^{(1)}$	$g_f V_4^8$		$F_1^{(1)}$	$(9)_{\alpha}^{(1)}$	$\frac{1}{\sqrt{2}} \left\{ -g_e \left(\frac{\sqrt{13}}{4} V_2^8 + \frac{\sqrt{3}}{4} V_6^8 \right) + g_f \left(\frac{\sqrt{7}}{2\sqrt{6}} V_4^8 + \frac{\sqrt{17}}{2\sqrt{6}} V_8^8 \right) \right\}$
	$F_1^{(2)}$	$(8)_{\alpha}^{(2)}$	$g_f V_8^8$		$F_1^{(2)}$	$(9)_{\alpha}^{(2)}$	$g_f U_2^8$
	$F_1^{(3)}$	$(8)_{\alpha}^{(3)}$	$g_e V_2^8$		$F_1^{(3)}$	$(9)_{\alpha}^{(3)}$	$g_f U_6^8$
	$F_1^{(4)}$	$(8)_{\alpha}^{(4)}$	$g_e V_6^8$		$F_1^{(4)}$	$(9)_{\alpha}^{(4)}$	$g_e W^8$
	$F_2^{(1)}$	$(8)_{\alpha}^{(1)}$	$g_e V_4^8$		$F_2^{(1)}$	$(9)_{\alpha}^{(1)}$	$g_e U_4^8$
	$F_2^{(2)}$	$(8)_{\alpha}^{(2)}$	$g_e V_8^8$		$F_2^{(2)}$	$(9)_{\alpha}^{(2)}$	$g_e U_8^8$
7	$F_2^{(3)}$	$(8)_{\alpha}^{(3)}$	$g_e V_2^8$	10	$F_2^{(3)}$	$(9)_{\alpha}^{(3)}$	$g_e U_2^8$
	$F_2^{(4)}$	$(8)_{\alpha}^{(4)}$	$g_e V_6^8$		$F_2^{(4)}$	$(9)_{\alpha}^{(4)}$	$g_e U_6^8$
	$F_2^{(5)}$	$(8)_{\alpha}^{(5)}$	$g_e V_8^8$		$F_2^{(5)}$	$(9)_{\alpha}^{(5)}$	$g_f W^8$
	$F_2^{(6)}$	$(8)_{\alpha}^{(6)}$	$g_e V_4^8$		$F_2^{(6)}$	$(9)_{\alpha}^{(6)}$	$g_f U_4^8$
	$F_2^{(7)}$	$(8)_{\alpha}^{(7)}$	$g_e V_8^8$		$F_2^{(7)}$	$(9)_{\alpha}^{(7)}$	$g_f U_8^8$
	$F_2^{(8)}$	$(8)_{\alpha}^{(8)}$	$g_e V_2^8$		$F_2^{(8)}$	$(9)_{\alpha}^{(8)}$	
	$F_2^{(9)}$	$(8)_{\alpha}^{(9)}$	$g_e V_6^8$		$F_2^{(9)}$	$(9)_{\alpha}^{(9)}$	
	$F_2^{(10)}$	$(8)_{\alpha}^{(10)}$	$g_e V_8^8$		$F_2^{(10)}$	$(9)_{\alpha}^{(10)}$	

10	$A_1^{(1)}$	$(10)_e^{(1)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{7\sqrt{15}}{8\sqrt{(2.29)}} U_2^{10} + \frac{3\sqrt{(13.15)}}{16\sqrt{29}} U_6^{10} + \frac{\sqrt{(13.17.19)}}{16\sqrt{29}} U_{10}^{10} \right) \right.$ $\left. + g_f \left(-\frac{\sqrt{(5.13)}}{4\sqrt{(2.29)}} U_4^{10} - \frac{\sqrt{(5.13.17)}}{8\sqrt{(6.29)}} U_8^{10} - \frac{\sqrt{(11.29)}}{8\sqrt{6}} W^{10} \right) \right\}$	10	$E^{(3)}$	$(10)_e^{(3)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\frac{\sqrt{(5.13)}}{4\sqrt{(2.29)}} U_4^{10} - \frac{\sqrt{(5.13.17)}}{8\sqrt{(6.29)}} U_8^{10} - \frac{\sqrt{(11.29)}}{8\sqrt{6}} W^{10} \right) \right.$ $\left. + g_f \left(\frac{7\sqrt{15}}{8\sqrt{(2.29)}} U_2^{10} + \frac{3\sqrt{(13.15)}}{16\sqrt{29}} U_6^{10} + \frac{\sqrt{(13.17.19)}}{16\sqrt{29}} U_{10}^{10} \right) \right\}$
	$A_1^{(2)}$	$(10)_e^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(\frac{\sqrt{(13.17)}}{4\sqrt{(3.29)}} U_2^{10} - \frac{11\sqrt{17}}{4\sqrt{(6.29)}} U_6^{10} + \frac{\sqrt{(5.19)}}{4\sqrt{(2.29)}} U_{10}^{10} \right) \right.$ $\left. + g_f \left(-\sqrt{\frac{17}{29}} U_4^{10} + \frac{2\sqrt{3}}{\sqrt{29}} U_8^{10} \right) \right\}$		$E^{(4)}$	$(10)_e^{(4)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\sqrt{\frac{17}{29}} U_4^{10} + \frac{2\sqrt{3}}{\sqrt{29}} U_8^{10} \right) \right.$ $\left. + g_f \left(\frac{\sqrt{(13.17)}}{4\sqrt{(3.29)}} U_2^{10} - \frac{11\sqrt{17}}{4\sqrt{(6.29)}} U_6^{10} + \frac{\sqrt{(5.19)}}{4\sqrt{(2.29)}} U_{10}^{10} \right) \right\}$
	$A_2^{(1)}$	$(10)_e^{(1)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\frac{\sqrt{(5.13)}}{4\sqrt{(2.29)}} U_4^{10} - \frac{\sqrt{(5.13.17)}}{8\sqrt{(6.29)}} U_8^{10} - \frac{\sqrt{(11.29)}}{8\sqrt{6}} W^{10} \right) \right.$ $\left. - g_f \left(\frac{7\sqrt{15}}{8\sqrt{(2.29)}} U_2^{10} + \frac{3\sqrt{(13.15)}}{16\sqrt{29}} U_6^{10} + \frac{\sqrt{(13.17.19)}}{16\sqrt{29}} U_{10}^{10} \right) \right\}$		$F_1^{(1)}$	$(10)_e^{(1)}$	$g_f V_4^{10}$
	$A_2^{(2)}$	$(10)_e^{(2)}$	$\frac{1}{\sqrt{2}} \left\{ g_e \left(-\sqrt{\frac{17}{29}} U_4^{10} + \frac{2\sqrt{3}}{\sqrt{29}} U_8^{10} \right) \right.$ $\left. - g_f \left(\frac{\sqrt{(13.17)}}{4\sqrt{(3.29)}} U_2^{10} - \frac{11\sqrt{17}}{4\sqrt{(6.29)}} U_6^{10} + \frac{\sqrt{(5.19)}}{4\sqrt{(2.29)}} U_{10}^{10} \right) \right\}$		$F_1^{(2)}$	$(10)_e^{(2)}$	$g_f V_8^{10}$
					$F_1^{(3)}$	$(10)_e^{(3)}$	$g_e V_2^{10}$
					$F_1^{(4)}$	$(10)_e^{(4)}$	$g_e V_6^{10}$
					$F_1^{(5)}$	$(10)_e^{(5)}$	$g_e V_{10}^{10}$
					$F_2^{(1)}$	$(10)_e^{(1)}$	$g_e V_4^{10}$
					$F_2^{(2)}$	$(10)_e^{(2)}$	$g_e V_8^{10}$
	$E^{(1)}$	$(10)_e^{(1)}$	$g_e \left(-\frac{\sqrt{11}}{4\sqrt{2}} U_4^{10} - \frac{\sqrt{(11.17)}}{8\sqrt{6}} U_8^{10} + \frac{\sqrt{(5.13)}}{8\sqrt{6}} W^{10} \right)$		$F_2^{(3)}$	$(10)_e^{(3)}$	$g_f V_{10}^{10}$
	$E^{(2)}$	$(10)_e^{(2)}$	$g_f \left(\frac{\sqrt{(13.19)}}{8\sqrt{6}} U_2^{10} + \frac{\sqrt{19}}{16\sqrt{3}} U_6^{10} - \frac{\sqrt{(5.17)}}{16} U_{10}^{10} \right)$		$F_2^{(4)}$	$(10)_e^{(4)}$	$g_f V_6^{10}$
					$F_2^{(5)}$	$(10)_e^{(5)}$	$g_f V_{10}^{10}$

TABLE IV. MATRIX ELEMENTS OF THE CORIOLIS OPERATOR A

R branch (J_{J-1}) levels				Q branch (J_J) levels				P branch (J_{J+1}) levels			
Level		Matrix element		Level		Matrix element		Level		Matrix element	
1_0	A_1	—	0	1_1	F_1	a	$\sqrt{2}$	0_1	F_2	—	0
2_1	F_2	a	$-\frac{3\sqrt{2}}{\sqrt{5}}$	2_2	E	a	$-2\sqrt{3}$	1_2	E	—	0
3	E	a	$\frac{2\sqrt{15}}{\sqrt{7}}$	3_3	A_2	—	0	2_3	A_1	a	$-2\sqrt{6}$
		a_1	$\frac{2\sqrt{15}}{\sqrt{7}}$			F_1	a_1			a	$-2\sqrt{3}$
		a_2	$\frac{2\sqrt{3}}{\sqrt{7}}$			F_2	a_2			a	$-\frac{2\sqrt{3}}{\sqrt{5}}$
4_3	A_1	a	$-2\sqrt{6}$	4_4	A_2	a	$-2\sqrt{14}$	3_4	A_1	—	0
		a_1	$-\sqrt{7}$			E	a_1			a	$\frac{6\sqrt{3}}{\sqrt{7}}$
		a_2	$2\sqrt{3}$			F_1	a_1			a_1	3
5_4	F_2	a_2	$-\sqrt{15}$	5_5	F_2	a_2	$-\sqrt{15}$	4_5	E	a_1	$-\frac{6\sqrt{3}}{\sqrt{7}}$
		a_1	$-\frac{6\sqrt{6}}{\sqrt{11}}$			a_1	$-\frac{14}{\sqrt{5}}$			a_2	$-\frac{2\sqrt{42}}{\sqrt{5}}$
		a_2	$-\frac{6\sqrt{42}}{\sqrt{55}}$			a_2	$-\frac{\sqrt{21}}{\sqrt{5}}$			a_1	$-\frac{2}{\sqrt{5}}$
	F_1	a_1	$\frac{6}{\sqrt{55}}$		$2F_1$	a_{11}	$-\sqrt{42}$		$2F_2$	a_{12}	$-2\sqrt{3}$
		a_3	$\frac{36}{\sqrt{55}}$			a_{12}	$\sqrt{30}$			a_{21}	$-\frac{12}{\sqrt{5}}$
		a_2	$\frac{6\sqrt{21}}{\sqrt{55}}$			a_{21}	$-\frac{3\sqrt{6}}{\sqrt{5}}$			a_{22}	$\frac{2\sqrt{21}}{\sqrt{5}}$
6_5	E	a_1	$-\frac{2\sqrt{(7.15)}}{\sqrt{13}}$	6_6	F_2	a_1	$\frac{3\sqrt{6}}{\sqrt{5}}$	5_6	A_1	a	$\frac{2\sqrt{210}}{\sqrt{11}}$
		a_2	$\frac{2\sqrt{33}}{\sqrt{13}}$			a_2	$-\sqrt{42}$			a	$2\sqrt{6}$
		a_3	$-\frac{2\sqrt{30}}{\sqrt{13}}$			a_3	$-\frac{3\sqrt{6}}{\sqrt{5}}$			E	a
	F_1	a_1	$-\frac{3\sqrt{30}}{\sqrt{13}}$		A_1	a	$\frac{2\sqrt{(3.55)}}{\sqrt{7}}$		F_1	a_1	$\frac{3\sqrt{30}}{\sqrt{11}}$
		a_2	$\frac{8\sqrt{3}}{\sqrt{13}}$			a	$-2\sqrt{3}$			a_3	$-\frac{4\sqrt{30}}{\sqrt{11}}$
		a_3	$-\frac{2\sqrt{30}}{\sqrt{13}}$			A_2	a				

TABLE IV (continued)

R branch (J_{J-1}) levels				Q branch (J_J) levels				P branch (J_{J+1}) levels					
Level		Matrix element		Level		Matrix element		Level		Matrix element			
6_5	$2F_2$	a_{12}	$\frac{2\sqrt{(7.15)}}{\sqrt{13}}$	6_6	F	a_1	$-\frac{\sqrt{6}}{\sqrt{330}}$	5_6	$2F_2$	a_{11}	$-\frac{8\sqrt{3}}{\sqrt{11}}$		
		a_{21}	$\frac{4\sqrt{30}}{\sqrt{13}}$			a_2	$\frac{\sqrt{7}}{12\sqrt{3}}$			a_{12}	$\frac{2\sqrt{(7.15)}}{\sqrt{11}}$		
		a_{22}	$\frac{3\sqrt{3}}{\sqrt{13}}$			a_3	$\frac{\sqrt{7}}{12\sqrt{3}}$			a_{13}	$-\frac{3\sqrt{3}}{\sqrt{11}}$		
		a_{23}	$-\frac{\sqrt{(11.15)}}{\sqrt{13}}$			$2F_1$	a_{11}			$-\frac{3\sqrt{15}}{\sqrt{7}}$	a_{23}	$\sqrt{15}$	
							a_{12}			$\frac{5\sqrt{6}}{\sqrt{7}}$			
		a_{21}	$-\frac{\sqrt{33}}{\sqrt{7}}$										
7_6	A_1	a	$\frac{6\sqrt{11}}{\sqrt{5}}$	7_7	A_2	a	$-\frac{2\sqrt{(11.13)}}{\sqrt{7}}$	6_7	A_1	a	$-\frac{12\sqrt{3}}{\sqrt{7}}$		
		A_2	a			$-\frac{18}{\sqrt{7}}$							
	E	a_1	$\frac{6\sqrt{26}}{\sqrt{35}}$			F_2	a_1			$-\frac{\sqrt{6}}{\sqrt{7}}$	E	a_1	$-\frac{3\sqrt{66}}{\sqrt{13}}$
		a_2	$\frac{6\sqrt{11}}{\sqrt{35}}$				a_2			$-\frac{3\sqrt{15}}{\sqrt{7}}$		a_2	$-\frac{9\sqrt{30}}{\sqrt{(7.13)}}$
	F_1	a_1	$\frac{5\sqrt{6}}{\sqrt{35}}$				a_3			$-\frac{\sqrt{33}}{\sqrt{7}}$		a_3	$-\frac{6\sqrt{33}}{\sqrt{(7.13)}}$
		a_2	$-\frac{\sqrt{(13.66)}}{\sqrt{35}}$	7_7	A_2	a	$-\frac{2\sqrt{(11.13)}}{\sqrt{7}}$	$2F_1$	a_{11}	$-\frac{15\sqrt{2}}{\sqrt{(7.13)}}$			
		a_4	$\frac{12\sqrt{11}}{\sqrt{35}}$			a	$-\frac{2\sqrt{(11.13)}}{\sqrt{7}}$		a_{12}	$-\frac{18\sqrt{5}}{\sqrt{(7.13)}}$			
	$2F_2$	a_{11}	$\frac{6\sqrt{3}}{\sqrt{7}}$		E	a_1	$-\frac{\sqrt{(13.22)}}{\sqrt{7}}$		a_{21}	$\frac{3\sqrt{22}}{\sqrt{7}}$			
		a_{13}	$\frac{2\sqrt{6}}{\sqrt{7}}$			a_2	$-\frac{4}{\sqrt{7}}$		a_{23}	$-\frac{18}{\sqrt{7}}$			
		a_{14}	$-\frac{3\sqrt{22}}{\sqrt{7}}$	$2F_1$	a_{11}	-9	$2F_2$	a_{12}	$-\frac{6\sqrt{10}}{\sqrt{13}}$				
		a_{22}	$\frac{18\sqrt{13}}{\sqrt{(7.15)}}$		a_{13}	$\sqrt{56}$		a_{21}	$-\frac{12\sqrt{33}}{\sqrt{(7.13)}}$				
		a_{24}	$\frac{3\sqrt{2}}{\sqrt{35}}$		a_{21}	$-\frac{5\sqrt{33}}{2\sqrt{7}}$		a_{22}	$\frac{3\sqrt{330}}{\sqrt{(7.13)}}$				
		a_{22}	$-\frac{3\sqrt{39}}{2\sqrt{7}}$		a_{23}	$-\frac{3\sqrt{6}}{\sqrt{(7.13)}}$							
8_7	A_1	a_1	$-\frac{12\sqrt{(13.21)}}{\sqrt{(17.31)}}$		8_8	A_1	a	$\frac{2\sqrt{21}}{\sqrt{5}}$	7_8	A_1	a	$\frac{2\sqrt{21}}{\sqrt{5}}$	
		a_2	$\frac{6\sqrt{55}}{\sqrt{(17.31)}}$										
	E	a_1	$-\frac{3\sqrt{14}}{\sqrt{17}}$	$2F_2$			a_{11}	$\frac{11\sqrt{2}}{\sqrt{7}}$			$2E$	a_{11}	$\frac{6\sqrt{(13.14)}}{\sqrt{(5.31)}}$
		a_2	$-\frac{6\sqrt{(11.21)}}{\sqrt{(17.31)}}$				a_{13}	-9				a_{12}	$\frac{6\sqrt{77}}{\sqrt{(5.31)}}$
		a_3	$\frac{24\sqrt{(5.13)}}{\sqrt{(17.31)}}$				a_{14}	$-\frac{5\sqrt{33}}{2\sqrt{7}}$				a_{21}	$-\frac{\sqrt{66}}{\sqrt{31}}$
					a_{22}	$-\frac{13\sqrt{2}}{\sqrt{7}}$	a_{22}	$-\frac{8\sqrt{39}}{\sqrt{31}}$					
					a_{24}	$-\frac{3\sqrt{39}}{2\sqrt{7}}$							

TABLE IV (continued)

<i>R</i> branch (<i>J_{J-1}</i>) levels					<i>Q</i> branch (<i>J_J</i>) levels																		
Level		Matrix element		Level		Matrix element		Level		Matrix element													
9 ₈	2 <i>F</i> ₂	<i>a</i> ₁₁	$\frac{2\sqrt{77}}{\sqrt{19}}$	10 ₉	2 <i>F</i> ₁	<i>a</i> ₁₁	$-\frac{\sqrt{(13.21)}}{\sqrt{5}}$	9 ₉	3 <i>F</i> ₁	<i>a</i> ₁₁	$-2\sqrt{33}$												
			<i>a</i> ₁₃				$\frac{2\sqrt{210}}{\sqrt{19}}$				<i>a</i> ₁₃	$\frac{24}{\sqrt{35}}$	<i>a</i> ₁₃	$3\sqrt{10}$									
							<i>a</i> ₁₄					$-\frac{2\sqrt{(13.33)}}{\sqrt{(2.19)}}$		<i>a</i> ₂₁	$\frac{3\sqrt{5}}{\sqrt{7}}$	<i>a</i> ₂₁	$-\frac{\sqrt{(13.21)}}{\sqrt{5}}$						
												<i>a</i> ₂₂			$\frac{18\sqrt{5}}{\sqrt{19}}$		<i>a</i> ₂₂	$\frac{12\sqrt{(3.17)}}{2\sqrt{35}}$	<i>a</i> ₂₂	$-\sqrt{35}$			
															<i>a</i> ₂₄			$\frac{\sqrt{70}}{\sqrt{19}}$		<i>a</i> ₂₄	$\frac{7\sqrt{2}}{\sqrt{5}}$	<i>a</i> ₂₄	$-\frac{2\sqrt{17}}{\sqrt{5}}$
																		<i>a</i> ₂₅			$-\frac{2\sqrt{170}}{\sqrt{19}}$		<i>a</i> ₂₅
10 ₉	<i>A</i> ₁	<i>a</i> ₁	$-\frac{72}{\sqrt{(7.29)}}$	3 <i>F</i> ₂	<i>a</i> ₁₃	$-\frac{3\sqrt{66}}{\sqrt{7}}$	2 <i>F</i> ₂	<i>a</i> ₁₁	$\frac{13\sqrt{2}}{\sqrt{5}}$														
			<i>a</i> ₂			$\frac{18\sqrt{(17.26)}}{\sqrt{(29.35)}}$			<i>a</i> ₂₁	$\frac{12\sqrt{2}}{\sqrt{5}}$	<i>a</i> ₁₃	$-2\sqrt{33}$											
	<i>A</i> ₂	<i>a</i> ₂			$-\frac{2\sqrt{(6.29)}}{\sqrt{5}}$	<i>a</i> ₂₃		$\frac{\sqrt{(6.13)}}{\sqrt{5}}$		<i>a</i> ₁₄		$-\frac{\sqrt{(13.21)}}{\sqrt{5}}$											
			<i>E</i>		<i>a</i> ₁			$-\frac{6\sqrt{66}}{\sqrt{35}}$	<i>a</i> ₂₄		$-2\sqrt{15}$	<i>a</i> ₂₂	$-\frac{3\sqrt{2}}{\sqrt{5}}$										
	$\frac{2\sqrt{(6.19)}}{\sqrt{35}}$	<i>a</i> ₃₂				$\frac{24\sqrt{6}}{\sqrt{35}}$		<i>a</i> ₂₄		$-\sqrt{35}$													
	$-\frac{12\sqrt{39}}{\sqrt{(7.29)}}$					<i>a</i> ₃₄				$\frac{\sqrt{(3.17)}}{\sqrt{35}}$	<i>a</i> ₂₅		$-\frac{2\sqrt{17}}{\sqrt{5}}$										
	$\frac{2\sqrt{(17.30)}}{\sqrt{(7.29)}}$									<i>a</i> ₃₅			$-\frac{3\sqrt{19}}{\sqrt{7}}$										

which gives us the perturbations of the two levels of ν_4 . Now it is easy to verify that these approximate roots are also the two roots of the following determinant

$$\begin{vmatrix} -\Delta x - (c_{11}^2 + c_{12}^2) & -(c_{11}c_{21} + c_{12}c_{22}) \\ -(c_{11}c_{21} + c_{12}c_{22}) & -\Delta x - (c_{21}^2 + c_{22}^2) \end{vmatrix} = 0,$$

i.e. the eigenvalues of the Hermitean square of the matrix of the coupling coefficients

$$CC' = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} = \begin{pmatrix} c_{11}^2 + c_{12}^2 & c_{11}c_{21} + c_{12}c_{22} \\ c_{21}c_{11} + c_{22}c_{12} & c_{21}^2 + c_{22}^2 \end{pmatrix}.$$

This is a general result holding for any degree of degeneracy (cf. E. Wigner 1931, p. 50): the approximations to the perturbed levels are found by diagonalization of the Hermitean square CC' of the matrix C of the coupling coefficients. In the general case

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \vdots & & & \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix},$$

$$CC' = \begin{pmatrix} c_{11}^2 + c_{12}^2 + \dots + c_{1n}^2, & \dots, & c_{11}c_{m1} + \dots + c_{1n}c_{mn} \\ \vdots & & \\ c_{m1}c_{11} + c_{m2}c_{12} + \dots + c_{mn}c_{1n}, & \dots, & c_{m1}^2 + c_{m2}^2 + \dots + c_{mn}^2 \end{pmatrix},$$

and if $\epsilon_1^2, \epsilon_2^2, \dots, \epsilon_m^2$ are the eigenvalues of this square of the coupling matrix, then the approximations to the perturbations are

$$x_1 = -\frac{\epsilon_1^2}{\Delta}, \quad x_2 = -\frac{\epsilon_2^2}{\Delta}, \quad \dots, \quad x_m = -\frac{\epsilon_m^2}{\Delta},$$

where $\Delta = E_2 - E_4$. We have used these approximate values since they are sufficient for our purpose, but they could of course be improved upon by direct diagonalization of the original energy matrix. Since in the levels up to $J = 10$ never more than three representations of the same type occur in the rotational sublevels of ν_4 , we have not had occasion to solve equations of higher order than the third.

In calculating the perturbation we must take into account the fact that owing to Teller's Coriolis coupling term $H^{(4)}$ the energy difference Δ between the pairs of levels is constant only for the Q branch (J_J) levels and varies

with J for the R and P branch (J_{J-1}, J_{J+1}) levels. Thus the unperturbed levels of ν_2 are given by

$$(J) = \nu_2 + B'_2 J(J+1),$$

and the unperturbed levels of ν_4 by

$$(R \text{ branch}) \quad (J_{J-1}) = \nu_4 + B'_4 J(J+1) - 2J\zeta_4 B'_4,$$

$$(Q \text{ branch}) \quad (J_J) = \nu_4 + B'_4 J(J+1),$$

$$(P \text{ branch}) \quad (J_{J+1}) = \nu_4 + B'_4 J(J+1) + 2(J+1)\zeta_4 B'_4.$$

We neglect any vibrational change in the moment of inertia, i.e. we put $B'_2 = B'_4 = B$. Then if Δ denotes the frequency difference $\nu_2 - \nu_4$, we must multiply the calculated perturbations $-\epsilon^2/\Delta$ by the factor $\frac{\nu_2 - \nu_4}{\nu_2 - \nu_4 + 2J\zeta B}$

for the J_{J-1} levels and by the factor $\frac{\nu_2 - \nu_4}{\nu_2 - \nu_4 - 2(J+1)\zeta B}$ for the J_{J+1} levels.

We see that this correcting factor reduces the perturbation in the R branch and increases it in the P branch. Since we are using only approximations to the true modes of vibration and further since the frequency ν_2 is not known exactly, we have taken a rounded off value of $-\frac{1}{20}$ for the factor with which to multiply the squares of the matrix elements of A to obtain the uncorrected perturbation

$$-\frac{|H^{(24)}|^2}{\nu_2 - \nu_4} = -\frac{11.0}{230} |A|^2 \sim -\frac{1}{20} |A|^2 \text{ cm.}^{-1}$$

In calculating the corrections to the P and R branch levels we have taken $B = 5.27$, $\zeta_4 = 0.45$.

The results of the perturbation calculation carried out in this manner are shown in Table V, both the corrected and uncorrected perturbations being given, in cm.^{-1} . We see that the correction is important. The sublevels of the individual rotational levels of ν_4 as here calculated are plotted on a common wave number scale in fig. 1. We notice from this diagram that the pattern of perturbed levels is similar (although different in magnitude) for the P and R branch levels, whilst for the Q branch levels the pattern is just reversed. This fact gives us a useful independent check on the calculations. We also see that the perturbation is greatest for the Q branch (J_J) levels, where for $J = 9$ some of the sublevels are perturbed by as much as 13 cm.^{-1}

TABLE V. PERTURBATION OF THE ROTATIONAL LEVELS OF ν_4

P branch (J_{J+1}) levels			Q branch (J_J) levels		R branch (J_{J-1}) levels		
Level	Perturbation (cm. ⁻¹)		Level	Perturbation (cm. ⁻¹)	Level	Perturbation (cm. ⁻¹)	
	Un-corrected	Corrected				Un-corrected	Corrected
$0_1 F_2$	0	0	$1_1 F_1$	-0.10	$1_0 A_1$	0	0
$1_2 E$	0	0	$2_2 F_1$	-0.30	$2_1 F_2$	-0.18	-0.17
$1_2 F_2$	-0.30	-0.31	$2_2 E$	-0.60	$3_2 E$	-0.43	-0.40
$2_2 F_2$	-0.12	-0.13	$3_2 F_2$	-0.75	$3_2 F_2$	-0.52	-0.49
$2_2 F_1$	-0.60	-0.64	$3_2 F_1$	-1.35	$4_2 F_2$	-0.75	-0.69
$2_2 A_1$	-1.20	-1.28	$4_2 F_1$	-0.37	$4_2 F_1$	-0.95	-0.88
$3_4 A_1$	0	0	$4_2 A_2$	-2.80	$4_2 A_1$	-1.20	-1.11
$3_4 F_1$	-0.45	-0.49	$5_2 F_1^{(1)}$	-0.69	$5_4 A_1$	-0.98	-0.89
$3_4 E$	-0.77	-0.84	$5_2 E$	-1.08	$5_4 F_1$	-1.21	-1.10
$3_4 F_2$	-1.74	-1.90	$5_2 F_2$	-3.18	$5_4 E$	-1.37	-1.24
$4_2 F_2^{(1)}$	-0.34	-0.38	$5_2 F_1^{(2)}$	-3.99	$5_4 F_2$	-1.87	-1.69
$4_2 F_1$	-0.88	-0.98	$6_2 A_2$	-0.60	$6_2 F_2^{(1)}$	-1.48	-1.32
$4_2 E$	-2.28	-2.54	$6_2 F_2$	-1.24	$6_2 F_1$	-1.78	-1.58
$4_2 F_2^{(2)}$	-2.54	-2.83	$6_2 F_1^{(1)}$	-1.97	$6_2 E$	-2.58	-2.30
$5_2 E$	-0.55	-0.63	$6_2 A_1$	-4.71	$6_2 F_2^{(2)}$	-2.72	-2.42
$5_2 F_2^{(1)}$	-0.71	-0.81	$6_2 F_1^{(2)}$	-5.49	$7_2 E$	-1.90	-1.66
$5_2 A_2$	-1.20	-1.37	$6_2 E$	-5.74	$7_2 F_2^{(1)}$	-2.00	-1.75
$5_2 F_2^{(2)}$	-2.95	-3.37	$7_2 F_1^{(1)}$	-1.12	$7_2 F_1^{(1)}$	-2.31	-2.02
$5_2 F_1$	-3.41	-3.89	$7_2 E$	-2.16	$7_2 A_2$	-3.41	-2.98
$5_2 A_1$	-3.82	-4.36	$7_2 F_2^{(1)}$	-2.83	$7_2 F_2^{(2)}$	-3.70	-3.23
$6_2 F_2^{(1)}$	-0.74	-0.86	$7_2 A_2$	-4.09	$7_2 F_1$	-3.96	-3.46
$6_2 F_1^{(1)}$	-1.01	-1.18	$7_2 F_2^{(2)}$	-7.46	$8_2 F_2^{(1)}$	-2.35	-2.02
$6_2 A_1$	-3.09	-3.61	$7_2 F_1^{(2)}$	-7.89	$8_2 F_1^{(1)}$	-2.53	-2.17
$6_2 F_1^{(2)}$	-3.86	-4.51	$8_2 F_1^{(1)}$	-1.52	$8_2 A_1$	-3.92	-3.36
$6_2 E$	-4.27	-4.99	$8_2 E^{(1)}$	-1.99	$8_2 F_1^{(2)}$	-4.43	-3.80
$6_2 F_2^{(2)}$	-4.91	-5.74	$8_2 F_2^{(1)}$	-3.56	$8_2 E$	-4.71	-4.04
$7_2 A_1$	-0.84	-1.01	$8_2 F_1^{(2)}$	-5.33	$8_2 F_2^{(2)}$	-5.14	-4.41
$7_2 F_1^{(1)}$	-1.02	-1.22	$8_2 E^{(2)}$	-10.11	$9_2 A_1$	-2.74	-2.31
$7_2 E^{(1)}$	-1.13	-1.35	$8_2 F_2^{(2)}$	-10.29	$9_2 F_1^{(1)}$	-2.87	-2.42
$7_2 F_2^{(1)}$	-4.00	-4.79	$8_2 A_2$	-10.60	$9_2 E^{(1)}$	-2.94	-2.48
$7_2 F_1^{(2)}$	-5.07	-6.07	$9_2 A_2$	-1.56	$9_2 F_2^{(1)}$	-4.95	-4.18
$7_2 E^{(2)}$	-6.01	-7.20	$9_2 F_2^{(1)}$	-2.16	$9_2 F_1^{(2)}$	-5.70	-4.81
$7_2 F_2^{(2)}$	-6.29	-7.53	$9_2 F_1^{(1)}$	-2.92	$9_2 E$	-6.36	-5.36
$8_2 F_2^{(1)}$	-1.11	-1.36	$9_2 A_1$	-4.76	$9_2 F_2^{(2)}$	-6.56	-5.53
$8_2 F_1^{(1)}$	-1.50	-1.84	$9_2 E$	-7.16	$10_2 F_2^{(1)}$	-3.26	-2.70
			$9_2 F_2^{(2)}$	-13.16	$10_2 F_1^{(1)}$	-3.36	-2.79
			$9_2 F_1^{(2)}$	-13.41	$10_2 E$	-5.93	-4.92
					$10_2 F_2^{(2)}$	-6.10	-5.06
					$10_2 A_2$	-6.96	-5.77
					$10_2 F_2^{(3)}$	-7.80	-6.47
					$10_2 F_1^{(2)}$	-8.08	-6.70
					$10_2 A_1$	-8.33	-6.91

CONCLUSION

We see from this calculation that the Coriolis coupling term $H^{(24)}$ has a very appreciable effect on the energy spectrum and in Part III, where we calculate the optical spectrum, we shall see that this splitting and displacement of the rotational levels is just what is required to explain the observed complex rotational structure of the infra-red absorption band of the fundamental vibration ν_4 .

SUMMARY

Using the vibrational-rotational wave functions derived in Part I, the perturbations of the rotational levels of the fundamental vibration ν_4 of the methane molecule, caused by the Coriolis term in the Hamiltonian which couples the rotational levels of ν_4 with the rotational levels of ν_2 , are calculated explicitly up to the tenth rotational quantum number. Table III, which gives the wave functions for the tetrahedral rotational sublevels of ν_2 , will be useful in calculating the perturbations of the rotational levels of any fundamental or overtone subvibrational level of symmetry type E in a tetrahedrally symmetrical molecule.

REFERENCES

- Jahn, H. A. 1935 *Ann. Phys.* **23**, 529.
— 1938 *Proc. Roy. Soc. A*, **168**, 469.
Tisza, L. 1933 *Z. Phys.* **82**, 48.
Wigner, E. 1931 "Gruppentheorie und Quantenmechanik." Braunschweig.
-

On the theory of coincidence experiments on cosmic rays

BY N. ARLEY

H. H. Wills Physical Laboratory, University of Bristol

(Communicated by N. F. Mott, F.R.S.)

INTRODUCTION

It is now generally agreed that the cosmic radiation consists of two groups called the soft and the hard component. This classification refers to the penetrating power of the particles in question. The soft component consists of ordinary electrons and light quanta, whereas it now seems very probable that the hard group consists of "heavy electrons". The phenomena connected with the soft component are described by the cascade theory of showers put forward by Bhabha and Heitler (1937) and Carlson and Oppenheimer (1937). It is the purpose of the present paper to work out the consequences of the theory in greater detail in order to make possible a more quantitative comparison between theory and experiment than has hitherto been possible.

This is desirable for several reasons: In the first place it is important to ascertain whether the phenomena connected with the soft component can be accounted for *completely* by the cascade theory or whether some of the facts require an explanation outside the present relativistic quantum mechanics.

In the second place it is of decisive importance to separate the effects due to the soft and hard component as far as possible in order to obtain a clear picture of the effects which are caused by the penetrating particles, the nature of which is now the main problem in cosmic-ray physics. Such a separation is of course only possible if the effects of the soft component are known to the greatest possible detail. It will be seen in the present paper that most of the standard counter experiments find a *quantitative* explanation in terms of the cascade theory. The experimental material available at present is not, however, sufficient for a complete discussion in the sense discussed above.

In this paper we calculate the probabilities of the different kinds of coincidences caused by the primary electrons. For this purpose it is first of all necessary to extend the calculations of Bhabha and Heitler, since, in

these calculations, only the number of electrons in a shower with energies larger than a certain critical energy is computed. In § 1 we calculate the number of electrons with energies smaller than this critical energy. In §§ 2–3 we then work out the mean probabilities of finding one, two or more particles in a shower, which are the quantities measured in the absorption curve and in the Rossi transition curve. We confine ourselves in this paper to the consideration of small showers, as the same discussion for bursts has been carried out by Euler (1938).

1. NUMBER OF SHOWER ELECTRONS BELOW THE CRITICAL ENERGY

In the cascade theory of Bhabha and Heitler ionization and the Compton effect were neglected, since all electrons in the shower were assumed to have energies larger than a certain critical energy. We shall now calculate the number of electrons with energies below this critical energy. The latter electrons are created partly by the faster electrons being slowed down by ionization, and partly by the Compton effect. The latter plays presumably an important role for the explanation of all backward effects (Geiger 1937). Since in these experiments the geometrical arrangements are rather difficult to take into account, we shall not attempt to treat them here. For all other experiments we believe the Compton effect to be of secondary importance compared with the effect of ionization. The probability of the former process is only large for comparatively much smaller energies, and both the energies and the number of the Compton electrons will therefore presumably be relatively small. We shall therefore in this paper neglect the Compton electrons and only take the effect of ionization into account.

This has already been done in the paper of Carlson and Oppenheimer, but their mathematical treatment is not very strict and it is difficult to obtain numerical results by their method.

In order to get some idea about the number of slower shower particles, we can proceed in the following way: We assume that above a certain critical energy, E_c , the electrons can only lose energy by radiation and below E_c only by ionization. For the critical energy E_c we take that energy at which the mean energy loss due to radiation is equal to that due to ionization. Bethe and Heitler (1934) have derived for E_c the approximate expression

$$E_c = \frac{1600mc^2}{Z}, \quad (1)$$

where mc^2 is the rest energy of the electron and Z the nuclear charge. As the correct expressions for the energy loss of an electron (Heitler 1936, fig. 22, p. 222) do not deviate considerably from such "step functions", the error due to this approximation will presumably not be very large. On this assumption the number of electrons which emerge after a certain thickness of material with energies above E_c has been worked out in the Bhabha-Heitler theory, and we now calculate the corresponding number below E_c .

In order that an electron should emerge with an energy below E_c it must once have had an energy greater than E_c and lost a light quantum with so high an energy that it remains with an energy less than E_c . We denote the probability that an electron with energy $E' (> E_c)$ should emit a light quantum with energy between k and $k + \Delta k$ while travelling the distance $\Delta l'$ by $P_{\text{rad}}(E', k) \Delta k \Delta l'$. The probability of losing an energy between E'' and $E'' + \Delta E''$ due to ionization in travelling the distance l'' we denote by $P_{\text{ion}}(E'', l'') \Delta E''$. We then obtain for the probability that an electron which at the distance l' had an energy E' will emerge at the distance l with an energy between E and $E + \Delta E$

$$\int_0^{E'} dk P_{\text{rad}}(E', k) \Delta l' P_{\text{ion}}(E' - k - E, l - l') \Delta E. \quad (2)$$

Multiplying (2) with the number of electrons and positrons which at the distance l' have an energy between E' and $E' + \Delta E'$, $p(l', E') \Delta E'$, which is given in the Bhabha-Heitler theory for $E' > E_c$, and integrating over E' from E_c to E_0 (E_0 is the energy of the incident electron producing the cascade) and over l' from 0 to l , we find that the total number of electrons and positrons which at the distance l emerge with an energy between $E (< E_c)$ and $E + \Delta E$ is

$$p^*(l, E) \Delta E = \int_0^l dl' \int_{E_c}^{E_0} dE' \int_0^{E'} dk P_{\text{rad}}(E', k) P_{\text{ion}}(E' - k - E, l - l') p(l', E') \Delta E. \quad (3)$$

Assuming now that the mean energy loss due to ionization is constant, i.e.

$$-\left(\frac{dE}{dl}\right)_{\text{ion}} = \beta, \quad (4)$$

which is correct down to energies of the order of 100,000 eV, our P_{ion} will be a δ -function

$$P_{\text{ion}}(E'', l'') \Delta E'' = \Delta(E'' \leq E_c) \delta(E'' - \beta l'') \Delta E'', \quad (5)$$

where the δ is a Dirac δ -function and the symbol Δ is defined by†

$$\Delta(E'' \leq E_c) = \begin{cases} 1 & \text{for } E \leq E_c, \\ 0 & \text{,, } E > E_c. \end{cases} \quad (6)$$

If $\Phi(E', k) \Delta k$ is the differential effective cross-section for the emission of a quantum with energy between k and $k + \Delta k$ by an electron with energy E' and σ is the number of nuclei per cm.³, then our $P_{\text{rad}}(E', k) \Delta k \Delta l'$ is given by

$$P_{\text{rad}}(E', k) \Delta k \Delta l' = \Delta(E' \geq E_c) \sigma \Phi(E', k) \Delta k \Delta l', \quad (7)$$

where Φ is given by (Bhabha and Heitler 1937)

$$\left. \begin{aligned} \Phi(E', k) \Delta k &= \bar{\Phi} F\left(\frac{k}{E'}\right) \frac{\Delta k}{k}, & F\left(\frac{k}{E'}\right) &\sim \frac{k}{E'} \frac{a}{[-\log(1 - k/E')]}, \\ \bar{\Phi} &= \frac{Z^2}{137} \left(\frac{e^2}{mc^2}\right)^2, & a \log 2 &= 4 \log(183 Z^{-1}) + \frac{2}{9}, \end{aligned} \right\} \quad (8)$$

the constant a being determined by the condition that the approximate F -function shall give the same mean energy loss as the correct F -function (Heitler 1936, eqn. 34, p. 172). For our purpose we can put

$$\Phi \Delta k = a \bar{\Phi} \delta\left(\frac{k}{E'}\right) \frac{\Delta k}{E'}, \quad (9)$$

since Φ will be essentially different from 0 only in the neighbourhood of $k/E' = 0$. Inserting the expressions (5), (7) and (9) into (3), we can at once perform all the integrations and we then obtain, if we now measure all lengths in units of $l_0 = \frac{1}{a\bar{\Phi}\sigma}$,‡

$$\left. \begin{aligned} p^*(l, E) \Delta E &= \Delta\left(l \geq \frac{E_c - E}{\beta}\right) p(L, E_c) \Delta E, \\ L &= l - \frac{E_c - E}{\beta}. \end{aligned} \right\} \quad (10)$$

For $E = E_c$ we have $L = l$ and therefore $p^*(E_c) = p(E_c)$. Thus the energy

† In this paper we shall denote by a symbol $\Delta(\)$ a function which is 1 when the condition stated in the brackets is fulfilled, and 0 elsewhere.

‡ This shower unit length has the value 0.358 cm. for Pb, 0.215 cm. for Au, 0.207 cm. for Pt, 0.824 cm. for Sn, 1.02 cm. for Cu, 1.26 cm. for Fe, 6.71 cm. for Al and 15.5 cm. for C in form of graphite, density 2.3 g./cm.³. (The value 7.8 for Al as given by Bhabha and Heitler was incorrect.)

distribution is continuous throughout as we should expect it to be, and has only a kink at the point $E = E_c$ due to the discontinuity of our P_{rad} and P_{ion} . This does not matter, however, for our further calculations, as we shall not use the energy spectrum itself but only the total average numbers, \bar{N}_f and \bar{N}_s of "fast" and "slow" electrons and positrons, denoting henceforth by "fast" electrons those with energies above E_c and by "slow" electrons those with energies below E_c .

The function $p(l, E)$ consists of two terms

$$p(l, E) = 2p_{\text{sec}}(l, E) + p_{\text{prim}}(l, E), \quad (11)$$

where p_{sec} is the energy distribution of the secondary positive or negative electrons in the cascade, and p_{prim} the probability distribution for the energy of the primary electron falling on the top of the plate. These two functions are given in the Bhabha-Heitler theory

$$\begin{aligned} p_{\text{sec}} &= -\frac{\partial f(l, y)}{\partial E} = \frac{e^y}{E_0} \frac{\partial f(l, y)}{\partial y}, \\ p_{\text{prim}} &= -\frac{\partial W(l, y)}{\partial E} = \frac{e^y}{E_0} \frac{e^{-y} y^{l-1}}{(l-1)!}, \\ y &= \log \frac{E_0}{E}, \end{aligned} \quad (12)$$

where $f(l, y)$ is the average number of electrons or positrons emerging at the distance l with energies greater than E , and $W(l, y)$ is the incomplete Γ -function, representing the probability of finding the primary electron at l with energy greater than E . Inserting (11) and (12) into (10) and using the fact that

$$\frac{\beta}{E_c} = \log 2, \quad (13)$$

which follows from the definition of E_c and from (4) and (8), we get taking $\epsilon = E/E_c$ as variable instead of E_c and dropping the asterisk

$$\begin{aligned} p_{\text{sec}}(l, \epsilon) &= \Delta \left(l \geq \frac{1-\epsilon}{\log 2} \right) \frac{\partial f(L, y_c)}{\partial y}, \\ p_{\text{prim}}(l, \epsilon) &= \Delta \left(l \geq \frac{1-\epsilon}{\log 2} \right) \frac{e^{-y_c} y_c^{L-1}}{(L-1)!}, \\ L &= l - \frac{1-\epsilon}{\log 2}, \quad \epsilon = \frac{E}{E_c}, \quad y_c = \log \frac{E_0}{E_c}. \end{aligned} \quad (14)$$

Finally, we obtain for the average total numbers $\bar{N}_s(\text{sec})$ and $\bar{N}_s(\text{prim})$

$$\left. \begin{aligned} \bar{N}_s(\text{sec}) &= 2 \int_0^1 p_{\text{sec}}(l, \epsilon) d\epsilon = 2 \log 2 \int_\gamma^1 \frac{\partial f(L, y_c)}{\partial y} dL, \\ \bar{N}_s(\text{prim}) &= \int_0^1 p_{\text{prim}}(l, \epsilon) d\epsilon = \log 2 \int_\gamma^1 \frac{e^{-y_c} y_c^{L-1}}{(L-1)!} dL, \\ \gamma &= \text{Max} \left(l - \frac{1}{\log 2}, 0 \right). \end{aligned} \right\} \quad (15)$$

We note that both expressions in (15) depend only on $y_c = \log E_0/E_c$. Using the figures for f and p given in Appendixes I and II, we have by graphical integration obtained the values for the two \bar{N}_s 's given in Table I where we also give the values for $\bar{N}_f(\text{sec}) = 2f(l, y_c)$ and for $\bar{N}_f(\text{prim}) = W(l, y_c)$.

In fig. 1 we give some of the spectra $p(l, E)$. As shown in Appendix II the p -function has as a function of l nearly the same form as the f -function, so that the shape of the spectrum for the slow electrons as a function of $\epsilon = E/E_c$ will be mainly determined by the shape of $f(l, y_c)$ as a function of l taken between the values $l - 1.44$ and l .

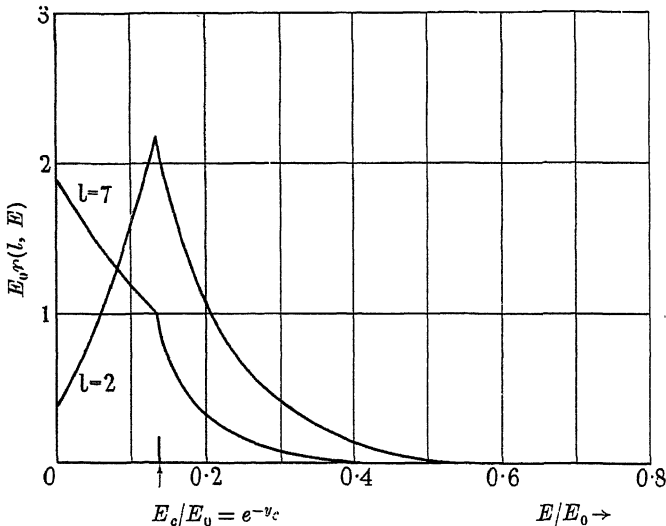


FIG. 1. The energy spectrum $E_0 p(l, E)$ for secondary electrons or positrons in a shower with $y_c = \log E_0/E_c = 2$.

From the physical meaning, as well as from the analytical expressions, it is easy to see that for negative values of y_c , i.e. for $E_0 \leq E_c$,† we have

† These values we shall need later for the calculation of the absorption curve (§ 5) since the primary electron, slowed down by ionization, will give a considerable contribution especially in light elements to this curve.

TABLE I. THE AVERAGE NUMBERS, \bar{N} , OF "FAST" (f) AND "SLOW" (s), PRIMARY (PRIM) AND SECONDARY (SEC) ELECTRONS AND POSITRONS AS FUNCTIONS OF l AND $y_c = \log E_0/E_c$. [\bar{N}_f (PRIM) = $W(l, y_c)$; \bar{N}_f (SEC) = $2 \times f(l, y_c)$, SEE APP. I]

	y_c	2	4	6	8	10	15
$l=0.2$	\bar{N}_f (prim)	0.987	1	1	1	1	1
	\bar{N}_f (sec)	0.0169	0.0462	0.078	0.110	0.143	0.226
	\bar{N}_s (sec)	0.0006	0.0007	0.001	0.001	0.001	0.001
	\bar{N}_s (prim)	0.0011	0.0001	0	0	0	0
$l=0.4$	\bar{N}_f (prim)	0.967	0.997	1	1	1	1
	\bar{N}_f (sec)	0.0602	0.171	0.298	0.426	0.550	0.900
	\bar{N}_s (sec)	0.0046	0.006	0.006	0.006	0.006	0.006
	\bar{N}_s (prim)	0.0050	0.001	0	0	0	0
$l=0.6$	\bar{N}_f (prim)	0.940	0.994	1	1	1	1
	\bar{N}_f (sec)	0.119	0.358	0.636	0.934	1.26	2.07
	\bar{N}_s (sec)	0.015	0.019	0.020	0.021	0.02	0.02
	\bar{N}_s (prim)	0.012	0.001	0	0	0	0
$l=0.8$	\bar{N}_f (prim)	0.906	0.989	1	1	1	1
	\bar{N}_f (sec)	0.185	0.590	1.09	1.64	2.23	3.94
	\bar{N}_s (sec)	0.033	0.044	0.05	0.05	0.05	0.06
	\bar{N}_s (prim)	0.024	0.002	0	0	0	0
$l=1$	\bar{N}_f (prim)	0.865	0.982	0.998	1	1	1
	\bar{N}_f (sec)	0.252	0.860	1.65	2.54	3.56	6.54
	\bar{N}_s (sec)	0.059	0.084	0.09	0.10	0.11	0.13
	\bar{N}_s (prim)	0.041	0.005	0	0	0	0
$l=2$	\bar{N}_f (prim)	0.594	0.908	0.983	0.997	1	1
	\bar{N}_f (sec)	0.494	2.44	5.80	10.7	17.3	43.4
	\bar{N}_s (sec)	0.332	0.62	0.89	1.1	1.3	2.1
	\bar{N}_s (prim)	0.174	0.04	0	0	0	0
$l=3$	\bar{N}_f (prim)	0.324	0.762	0.938	0.986	0.997	1
	\bar{N}_f (sec)	0.604*	3.90*	11.7*	25.8*	49.2*	155*
	\bar{N}_s (sec)	0.654	1.66	2.9	4.3	6.4	14
	\bar{N}_s (prim)	0.230	0.09	0	0	0	0
$l=5$	\bar{N}_f (prim)	0.053	0.371	0.715	0.900	0.971	1
	\bar{N}_f (sec)	0.448*	4.80*	25.5*	79.6*	205*	1260*
	\bar{N}_s (sec)	0.780	4.18	13.5	31.8	53	210
	\bar{N}_s (prim)	0.154	0.21	0.1	0	0	0
$l=10$	\bar{N}_f (prim)	0	0.008	0.084	0.284	0.543	0.930
	\bar{N}_f (sec)	0.0226	1.10*	20.8	180*	1050	28,000
	\bar{N}_s (sec)	0.0872	2.98	32.6	191	730	11,200
	\bar{N}_s (prim)	0.0006	0.02	0.1	0	0	0
\bar{N}_f (sec)	$l=14$	2.74×10^{-3}	0.24*	7.38	120*	1070	86,000
	$l=18$	3.22×10^{-4}	3.8×10^{-2} *	2.02	50*	668	142,000
	$l=24$	1.25×10^{-5}	2.8×10^{-3} *	0.198	8.2*	163	120,000
	$l=36$	1×10^{-8} *	1×10^{-5} *	2×10^{-3} *	0.1*	6*	14,500

* Interpolated.

$\bar{N}_j(\text{sec}) = \bar{N}_j(\text{prim}) = \bar{N}_s(\text{sec}) = 0$. The value of $\bar{N}_s(\text{prim})$ for $y_c < 0$ is obtained by an expression analogous to (3):

$$\bar{N}_s(\text{prim}) = \Delta\{\log(\log 2 \cdot l) \leq y_c \leq 0\}, \quad (16)$$

which only expresses the fact that if the plate is sufficiently thick, the primary particle is absorbed (i.e. if $\beta l > E_0$).

For the moment it is not possible to compare our results with experiments as no relevant data seem to be available. We must, however, bear in mind, as pointed out by Wilson,[†] that electrons with energies below E_c will have a rather great chance of being scattered in the material so that it is no longer legitimate to treat the problem as a one-dimensional problem. As a consequence of this scattering, fewer slow electrons will emerge from the plate so that we should expect the experimental values of \bar{N}_s to be smaller than the calculated ones. It is, however, probable that this effect would not alter the order of magnitude of \bar{N}_s by more than, say, a factor 2, nor will it influence the general shape of the spectrum considerably.

2. THE FLUCTUATION FORMULA

All the figures obtained so far are only *average* values and we must, therefore, discuss the fluctuations of the N 's about these average values before we can proceed further. Bhabha and Heitler (1937) assumed that in the shower the probability for a secondary electron to emerge with energy between E_1 and $E_1 + \Delta E_1$ is independent of the probability for another secondary electron to emerge with energy between E_2 and $E_2 + \Delta E_2$. It may, of course, be doubted whether this is really true as both electrons are generated in one multiplication process from the same initial parent electron, and, therefore, the two probabilities mentioned are in some way connected.

From this independence assumption, it follows at once that the probability $P(N)$ of finding N secondary particles when the average number is \bar{N} is given by the Poisson fluctuation formula

$$P(N) = \frac{e^{-\bar{N}} \bar{N}^N}{N!}. \quad (17)$$

The use of this fluctuation formula has been criticized from various sides, e.g. Furry (1937) and Euler (1938).[‡] To avoid the independence assumption,

[†] In a private discussion with the author.

[‡] Private communication; as Euler is, however, mainly interested in large showers with 100 or more particles, his considerations are not so relevant here.

Furry has proposed the following model for the multiplication process. He assumes that an electron has the probability Δl of creating a new electron by travelling a distance Δl and that the electron does not lose energy by this process. Thus neglecting entirely the important feature of the absorption of the electrons in the cascade, he obtains, therefore, the result that the average number of particles \bar{N} increases indefinitely with the thickness as e^l . Only for very small thicknesses are these assumptions justified, but, in this case, Furry's fluctuation formula deviates but little from the Poisson formula. For larger thicknesses it is certainly not legitimate to neglect *both* the energy loss *and* the light quanta. It seems, therefore, doubtful to us whether Furry's treatment of the problem can be regarded as a better approximation, since his idealization leaves only little resemblance with the real multiplication process.

In order to see whether it is legitimate or not to use the Poisson fluctuation formula for *small* showers, we have investigated the following model which, in this case (i.e. not too large values of the γ variable), is a rather close idealization of the real multiplication process. We assume firstly that an electron by travelling the distance Δl has the probability Δl of being absorbed by emitting two light quanta. (In reality an electron loses its energy in the form of a few quanta rather than in emitting one single quantum.) Secondly, we assume that a light quantum by travelling the distance Δl has the probability Δl of being absorbed creating two electrons. In order to account for the energy degradation taking place in each such step, we thirdly assume that when three such transformations have occurred, the resulting light quanta have lost so much energy that they can no longer create electrons and that no further multiplication takes place. The maximum number of particles in this cascade is, therefore, 4. Denoting by $P(n, m, l)$ the probability that n (≤ 4) electrons and m (≤ 8) light quanta are present at the distance l , and taking as initial condition $P(1, 0, 0) = 1$, all other P 's equal to 0 for $l = 0$, we can easily write down and solve the differential equations determining the ten possible P 's. Following the multiplication process step by step, we obtain

$$\frac{d}{dl} P(1, 0, l) = -P(1, 0, l), \quad \therefore P(1, 0, l) = e^{-l};$$

$$\frac{d}{dl} P(0, 2, l) = P(1, 0, l) - 2P(0, 2, l), \quad \therefore P(0, 2, l) = e^{-l} - e^{-2l};$$

$$\frac{d}{dl} P(2, 1, l) = 2P(0, 2, l) - 3P(2, 1, l), \quad \therefore P(2, 1, l) = e^{-l} - 2e^{-2l} + e^{-3l};$$

and so on. All the equations can be solved one after the other by elementary calculation. Denoting by $Q(N, l)$ the probability that N electrons emerge at the distance l and by \bar{N} the average number, we finally have

$$\left. \begin{aligned} Q(0, l) &= P(0, 2, l) + P(0, 5, l) + P(0, 8, l) \\ &= 1 + \left(\frac{7}{3} + 2l - 2l^2\right) e^{-l} + (-3 - 4l - 4l^2) e^{-2l} - 2le^{-3l} - \frac{1}{3}e^{-4l}, \\ Q(1, l) &= P(1, 0, l) + P(1, 3, l) + P(1, 6, l) \\ &= \left(-\frac{10}{3} - 4l + 2l^2\right) e^{-l} + (4 + 4l + 8l^2) e^{-2l} + (-1 + 6l) e^{-3l} + \frac{4}{3}e^{-4l}, \\ Q(2, l) &= P(2, 1, l) + P(2, 4, l) \\ &= (-1 + 2l) e^{-l} + (4l - 4l^2) e^{-2l} + (3 - 6l) e^{-3l} - 2e^{-4l}, \\ Q(3, l) &= P(3, 2, l) \\ &= \frac{5}{3}e^{-l} - 4le^{-2l} + (-3 + 2l) e^{-3l} + \frac{4}{3}e^{-4l}, \\ Q(4, l) &= P(4, 0, l) \\ &= \frac{1}{3}e^{-l} - e^{-2l} + e^{-3l} - \frac{1}{3}e^{-4l}; \\ \bar{N} &= \sum_{N=0}^4 NQ(N, l) = (1 + 2l^2) e^{-l}. \end{aligned} \right\} \quad (18)$$

According to Bhabha and Heitler (1937) the Poisson formula (17) was only derived for the *secondary* electrons and the *primary* electron had to be treated separately, having a probability $W(l, y)$ to be found at a thickness l . In our model the latter probability is simply $P(1, 0, l) = e^{-l}$. Thus the probability $P^*(N, l)$ of finding N secondary or primary electrons at the distance l is, therefore, according to (17), given by the modified Poisson formula

$$\left. \begin{aligned} P^*(N, l) &= e^{-\bar{N}_{\text{sec}}} \frac{\bar{N}_{\text{sec}}^{(N-1)}}{(N-1)!} \left[P(1, 0, l) + [1 - P(1, 0, l)] \frac{\bar{N}_{\text{sec}}}{N} \right], \\ \bar{N}_{\text{sec}} &= \sum_{N=0}^4 N[Q(N, l) - \delta_{N,1} P(1, 0, l)] = 2l^2 e^{-l}, \quad P(1, 0, l) = e^{-l}, \end{aligned} \right\} \quad (19)$$

where \bar{N}_{sec} is the mean number of secondary particles in our model. This has now to be compared with the exact expressions derived from our model. In Table II we give $Q(N, l)$ and $P^*(N, l)$ for a few values of l and we see that they agree very well with each other.

TABLE II. THE FLUCTUATION FORMULA DERIVED FROM OUR MODEL
COMPARED WITH THE MODIFIED POISSON FORMULA

	$Q(0, l)$	$P^*(0, l)$	$Q(1, l)$	$P^*(1, l)$	$Q(2, l)$	$P^*(2, l)$	$Q(3, l)$	$P^*(3, l)$	$Q(4, l)$	$P^*(4, l)$	$P^*(> 4, l)$	\bar{N}_{sec}
$l=0.5$	0.24	0.29	0.64	0.54	0.10	0.15	0.01	0.02	0.01	0	0	0.30
$l=2$	0.27	0.29	0.38	0.36	0.24	0.22	0.08	0.09	0.03	0.03	0	1.08
$l=3$	0.39	0.39	0.35	0.37	0.19	0.17	0.05	0.06	0.01	0.01	0	0.90

Our model could, of course, easily be extended to give larger values of \bar{N}_{sec} (the maximum value is here 1.1 for $l = 2$) by allowing for more steps before the multiplication stops. The calculation would still be elementary but more tedious, and we believe the result obtained would not be changed essentially even if \bar{N}_{sec} was increased by a factor five or ten. Only for large showers with 100 or more particles the Poisson formula is no longer valid, as is shown by the considerations of Euler (1938). We have, therefore, decided to take for our purpose the modified Poisson formula P^* as given in (19) with† $P(1, 0, l) \rightarrow W(l, y_c)$ and $\bar{N}_{\text{sec}} \rightarrow \bar{N} = 2f(l, y_c) + \bar{N}_s(l, y_c)$, as given in Table I. Denoting henceforth by $P(N, l, y_c)$ the probability for N particles, electrons plus positrons, to emerge at the distance l , if the energy of the incident electron (which we shall denote by E instead of E_0) is given by $\log E/E_c = y_c$, we therefore have for positive values of y_c

$$\left. \begin{aligned} P(0, l, y_c) &= e^{-\bar{N}}[1 - W(l, y_c)], \\ P(N, l, y_c) &= e^{-\bar{N}} \frac{\bar{N}^{N-1}}{(N-1)!} \left[W(l, y_c) + [1 - W(l, y_c)] \frac{\bar{N}}{N} \right], \\ P(> N, l, y_c) &= 1 - e^{-\bar{N}} \sum_{i=0}^N \frac{\bar{N}^i}{i!} + W(l, y_c) e^{-\bar{N}} \frac{\bar{N}^N}{N!}, \\ \bar{N} &= \bar{N}_f(\text{sec}) + \bar{N}_s(\text{sec}) + \bar{N}_s(\text{prim}), \\ y_c &> 0. \end{aligned} \right\} \quad (20)$$

If the primary energy is smaller than E_c , $y_c \leq 0$ (a case which will also be necessary to consider), (20) is no longer valid. As, however, the only particle capable of emerging from the plate in this case is the primary electron itself slowed down by ionization, we obtain at once from (16)

$$\left. \begin{aligned} P(0, l, y_c) &= \Delta\{y_c < \log(\log 2 \cdot l)\}, \\ P(1, l, y_c) &= \Delta\{\log(\log 2 \cdot l) \leq y_c \leq 0\}, \\ P(> 1, l, y_c) &= 0, \\ y_c &\leq 0. \end{aligned} \right\} \quad (21)$$

3. ENERGY SPECTRUM OF THE SOFT COMPONENT

What is measured in all coincidence experiments is the *mean* effect produced by all the incident electrons hitting the material dealt with in the

† Strictly speaking, it is not correct to use the Poisson formula for the distribution of the primary slow electron which is included in \bar{N}_s , but if $\bar{N} \ll 1$ the error involved in taking the Poisson formula $P(0) = e^{-\bar{N}}$, $P(1) = e^{-\bar{N}}\bar{N}$, $P(> 1) = 1 - e^{-\bar{N}}(1 + \bar{N})$, instead of the correct $P(0) = 1 - \bar{N}$, $P(1) = \bar{N}$, $P(> 1) = 0$, is very small.

experiment. In order to compare our theory with experiments we must, therefore, first consider the energy spectrum of the primary electrons producing the cascades. Next we must calculate the mean value of the expressions $P(N, l, y_c)$, given in (20) and (21), with respect to this energy spectrum

$$\left. \begin{aligned} \bar{P}(N, l) &= \int_0^\infty P(N, l, E) F(E) dE, \\ \int_0^\infty F(E) dE &= 1, \quad \therefore \sum_{N=0}^\infty \bar{P}(N) \equiv 1, \end{aligned} \right\} \quad (22)$$

where $F(E) \Delta E$ denotes the probability for a primary electron or positron to have an energy between E and $E + \Delta E$. For high energies this spectrum has been calculated from the intensity curve in the atmosphere by Heitler (1937*a*)

$$F(E) \Delta E \propto \frac{\Delta E}{E^{\gamma+1}} \quad \text{for } E > E_c^{\text{air}}, \quad (23)$$

where E_c^{air} is the critical energy for air (1.5×10^8 eV) and γ is a function of E which varies only slowly and can be put equal to $\frac{3}{2}$. This spectrum cannot be checked directly by experiments, since for high energies most of the particles observed belong to the hard component. Some indirect evidence can, however, be obtained by considering the ratio of the numbers of large and small bursts (Blackett 1938) and it is found that the spectrum agrees well with (23). The spectrum for energies below E_c^{air} could be obtained by integrating the number of slow electrons in air showers as given in § 1 over the spectrum of the primary electrons arriving at the top of the atmosphere. As at sea-level we are well beyond the maximum of the corresponding $f(l, y_c)$ functions, we would find that the spectrum below E_c^{air} increases slightly with decreasing energy from the value for $E = E_c^{\text{air}}$ or is practically constant (cf. § 1). The spectrum which is deduced from experiments has however quite a different form (Blackett 1938). Above a certain energy, $E_0 \sim 2 \times 10^8$ eV, which happens accidentally to be of the same magnitude as E_c^{air} , the spectrum is, as already mentioned, in agreement with (23). Below E_0 , however, the number of electrons found is much larger (perhaps 5–20 times at sea-level) than would correspond to the above calculations. This fact is explained by Blackett (1938) by assuming that the particles of the hard component transform themselves into ordinary electrons or produce secondary electrons when their energy passes below 2×10^8 eV.

The form of the spectrum below E_0 is not quite known, but we assume it to be roughly constant, since it will be seen that a deviation from the constant form is not very important (eqn. (25)). Consequently for the

spectrum $F(E) \Delta E$ of the primary electrons producing the cascades we have used the form

$$\left. \begin{aligned} F(E) \Delta E &= \frac{\alpha}{\alpha + 1/\gamma} \frac{\Delta E}{E_0} \quad \text{for } 0 \leq E < E_0, \\ F(E) \Delta E &= \frac{1}{\alpha + 1/\gamma} \left(\frac{E_0}{E} \right)^{\gamma+1} \frac{\Delta E}{E_0} \quad \text{for } E_0 \leq E, \\ \int_0^\infty F(E) dE &= 1, \end{aligned} \right\} \quad (24)$$

where α is the ratio between the constant value of F for $E < E_0$ and the value of F for $E = E_0$. If we now transform to the y_c variable, we have to multiply (24) with $dE/dy_c = E$ and finally we have, again denoting the spectrum by $F(y_c) \Delta y_c$,

$$\left. \begin{aligned} F(y_c) \Delta y_c &= \frac{\alpha}{\alpha + 1/\gamma} e^{-(y_c^0 - y_c)} \Delta y_c, \quad -\infty \leq y_c < y_c^0, \\ F(y_c) \Delta y_c &= \frac{1}{\alpha + 1/\gamma} e^{-\gamma(y_c - y_c^0)} \Delta y_c, \quad y_c^0 \leq y_c, \\ y_c^0 &= \log \frac{E_0}{E_c}, \quad \int_{-\infty}^{+\infty} F(y_c) dy_c = 1. \end{aligned} \right\} \quad (25)$$

It is seen that this function has a very sharp maximum at the point $y_c = y_c^0$, so that the mean values of the $P(N, l, y_c)$'s as given in (22) are practically determined by their behaviour in the neighbourhood of $y_c = y_c^0$. The results will, therefore, in most cases be seen to depend sensitively neither upon γ nor on α . Introducing (25) into (22), $\bar{P}(N, l)$ is given by

$$\begin{aligned} \bar{P}(N, l) &= \frac{\alpha}{\alpha + 1/\gamma} \int_{-\infty}^{y_c^0} e^{-(y_c^0 - y_c)} P(N, l, y_c) dy_c \\ &\quad + \frac{1}{\alpha + 1/\gamma} \int_{y_c^0}^{\infty} e^{-\gamma(y_c - y_c^0)} P(N, l, y_c) dy_c. \end{aligned} \quad (26)$$

For the numerical values we have taken $\gamma = \frac{3}{2}$ and $E_0 = 2 \times 10^8$ eV leaving the value of α for the present undetermined (experiments seem to give values of the order 5 to 20 at sea-level) since α is the most unknown factor in the spectrum. With this value for E_0 we have simply according to (1)

$$y_c^0 = \log \frac{E_0}{E_c} = \log \frac{Z}{4} = \begin{cases} 3.0 & \text{for Pb,} \\ 1.9 & \text{for Fe,} \\ 1.2 & \text{for Al.} \end{cases} \quad (27)$$

In Table III we give the values of the two integrals in (26), $\int_{<E_0}$ and $\int_{>E_0}$ for Pb, obtained by graphical integration. In Table IV the corresponding values are given for Al, and the values for intermediate elements will then lie between the lead and the aluminium values.

TABLE III. $\int_{-\infty}^{y_c^0} e^{-(y_c^0 - y_c)} P(N, l, y_c) dy_c = \int_{<E_0}$ AND
 $\int_{y_c^0}^{\infty} e^{-1.5(y_c - y_c^0)} P(N, l, y_c) dy_c = \int_{>E_0}$ FOR LEAD

$N \backslash l$	0.2	0.4	1	2	3	5	10
0 $\int_{<E_0}$	0.018	0.0381	0.108	0.209	0.223	0.276	0.679
$\int_{>E_0}$	~ 0	~ 0	0.0139	0.0069	0.0043	0.0032	0.0768
1 $\int_{<E_0}$	0.964	0.895	0.613	0.275	0.201	0.200	0.218
$\int_{>E_0}$	0.631	0.578	0.281	0.0643	0.0277	0.0149	0.127
2 $\int_{<E_0}$	0.0195	0.0679	0.232	0.250	0.207	0.179	0.0734
$\int_{>E_0}$	0.0267	0.0918	0.235	0.133	0.0619	0.0373	0.117
3 $\int_{<E_0}$	~ 0.0005	~ 0.002	0.0448	0.156	0.165	0.138	0.0219
$\int_{>E_0}$	~ 0.0015	~ 0.002	0.0928	0.158	0.0939	0.0555	0.0864
4 $\int_{<E_0}$	0	~ 0	0.00875	0.0618	0.109	0.0938	0.0042
$\int_{>E_0}$	0	~ 0	0.0289	0.129	0.105	0.0672	0.0619
$> 4 \int_{<E_0}$	0	0	0.00162	0.0483	0.0990	0.113	0.00096
$\int_{>E_0}$	0	~ 0	0.00918	0.177	0.389	0.480	0.200

TABLE IV. $\int_{-\infty}^{y_c^0} e^{-(y_c^0 - y_c)} P(N, l, y_c) dy_c = \int_{< E_0}$ AND
 $\int_{y_c^0}^{\infty} e^{-1.5(y_c - y_c^0)} P(N, l, y_c) dy_c = \int_{> E_0}$ FOR ALUMINIUM

$\begin{smallmatrix} l \\ N \end{smallmatrix}$	0.2	0.4	1	2	3	5	10
0 $\int_{< E_0}$	0.047	0.115	0.302	0.756	0.800	0.874	0.986
$\int_{> E_0}$	~ 0	~ 0.0015	0.0811	0.169	0.156	0.254	0.539
1 $\int_{< E_0}$	0.956	0.891	0.674	0.167	0.139	0.105	0.0124
$\int_{> E_0}$	0.635	0.604	0.434	0.229	0.204	0.197	0.0928
2 $\int_{< E_0}$	~ 0.002	~ 0.008	0.0267	0.0717	0.049	0.020	~ 0.001
$\int_{> E_0}$	~ 0.01	~ 0.02	0.111	0.171	0.144	0.106	0.0213
3 $\int_{< E_0}$	0	~ 0.0002	~ 0.001	0.00309	0.0107	~ 0	0
$\int_{> E_0}$	0	~ 0.003	0.0207	0.0587	0.0790	0.0512	0.0085
4 $\int_{< E_0}$	0	0	~ 0.0003	0.00155	0.0015	0	0
$\int_{> E_0}$	0	0	~ 0.002	0.0258	0.0395	0.0256	0.0043
> 4 $\int_{< E_0}$	0	0	0	~ 0	~ 0	0	0
$\int_{> E_0}$	0	0	~ 0.0002	0.0171	0.0459	0.0480	0.0139

4. GENERAL DISCUSSION OF COUNTER EXPERIMENTS

From the figures given in Tables III and IV it is now possible according to (26) to give a quantitative theory for many of the standard counter experiments. In discussing these experiments one has, however, to take into account the fact that in reality the phenomena are not one-dimensional, an assumption which is made throughout the whole cascade theory, but

that the different *geometrical factors* have to be considered. The Rossi curves, for instance, giving triple or higher coincidences as functions of the thickness of the shower producing materials, are determined firstly by the different probabilities, $\bar{P}(N, l)$, that N particles will emerge, and, secondly, by the probabilities, g_N , which we call the geometrical factors, of these N particles producing a coincidence. The geometrical factors depend on the angular spread in the showers, the geometrical arrangement of material and counters,* the sensitivity of the counters and so on. Since these factors are in most cases unknown,† we cannot expect more than a rough agreement with experiment, quite apart from the inaccuracies involved in the latter. From our tables it is possible to obtain the theoretical curves, $\sum_N g_N \bar{P}(N, l)$,

corresponding to arbitrary values for the geometrical factors. For reasons of simplicity, we have taken the non-vanishing g_N 's equal to unity throughout, and, as will be seen from the discussion in §§ 5 and 6, the agreement with experiments is as good as can be expected. It seems, therefore, hardly necessary to consider other values for the geometrical factors except perhaps for very special experimental arrangements (very large angles, etc.).

Furthermore, we must take into account the fact that the experiments as a rule give not only the showers due to the *electrons* and *positrons*, but also those due to *light quanta* and the *hard component*. The effect due to the *light quanta* could be subtracted by placing one or more counters above the shower producing plate. This has been done by Clay and van Gemert (1936) and by Maass (1936). The former measured triple coincidences from a lead plate 5×36 cm covered by six counters in parallel. The result was that the coincidence curves obtained with and without the six upper counters in series with the three lower counters were identical. This result seems at first sight surprising as it is known from both theory and Wilson chamber photographs that the number of light quanta in air is of the same order of magnitude as the number of electrons. As the energy spectrum of the light quanta is similar to that of the electrons and the showers created by light quanta are of the same type (except for very small thicknesses), we should expect that the two curves with and without the upper counters differ by a constant factor. This is contrary to what is found experimentally. It has, however, been pointed out by Clay himself‡ and by Blackett§

* The important role of the geometrical arrangement of the counters is clearly shown in the experiments of Zeiller (1935).

† A very interesting attempt to determine the geometrical factors experimentally has been described by Geiger (1937).

‡ In a conversation with Heitler.

§ In a conversation with the author.

that it is not at all improbable that the electrons and light quanta from air arrive in groups simultaneously and separated by only small distances, because all these particles are parts of a shower created in the atmosphere.* It is therefore likely that each time the shower producing plate is hit by a light quantum, it is also hit by an electron and vice versa. This would explain the result mentioned above.

In the experiments of Maass (1936) double coincidences are measured with different thicknesses of iron placed between the two counters (*A* curve) and above them (*B* curve), the area of the iron sheets being equal to the area of the cross-section of the counters. The curve *B-A* then gives the effect due to the non-ionizing radiation alone. Maass finds that the difference *B-A* is larger than the experimental errors but that the difference is small compared with *A* or *B*. From this experiment we therefore conclude that there is a certain but only small probability of a light quantum hitting the material without being accompanied by an electron. The experiments of Maass are, therefore, in agreement with those of Clay and with our interpretation.

Furthermore, we believe that the relatively large background always found in counter experiments also indicates that there is a comparatively large probability for two or more electrons arriving simultaneously at small distances (Hilgert and Bothe 1936; Schmeiser and Bothe 1938; the accidental counts due to the finite resolving power of the counting apparatus itself are certainly, as a rule, much smaller than the measured background). It seems, therefore, desirable to obtain more experimental evidence about this point; for instance, by measuring double coincidences in free air as a function of the distance between the counters without any material placed above them and so on†. For the reasons mentioned above, we shall in this paper assume that the presence of the showers produced by light quanta does not, except for small thicknesses, essentially influence the shape of the measured curves, but only their absolute magnitude.‡

The effect due to the *hard component* is as a rule subtracted, if at all, by extrapolating backwards from large thicknesses. Schwegler (1935) has

* According to the cascade theory we find, for instance, that the mean distance at sea-level between two electrons created at a height of 10 km. by an electron with energy $E \sim 10^{10}$ eV is only of the order 1 m., since the mean angle is of the order mc^2/E . Furthermore, we must remember that the hard component also creates showers in the atmosphere and at much lower heights so that the spread of these showers will probably be still smaller.

† [Note added in proof.] Such experiments have been performed by Auger and Maze (1938) and results are in accordance with the views held here.

‡ We intend to treat the showers due to the light quanta in another paper.

tried to subtract the effect produced by the hard component experimentally. He measured triple coincidences from a lead plate, P , using the arrangement shown in fig. 2. The curve I is obtained without the 10 cm. lead block, B , the curve II with B placed between the counters. The curve III, giving the difference between I and II is then thought of as giving the effects due to the soft component alone. This is, however, not quite certain since the hard component will produce showers in the lead block which may also emerge in the backward direction (Heitler 1938), thus producing coincidences with the two upper counters. If, on the other hand, the lead block had been placed above the shower producing plate, the showers produced in the

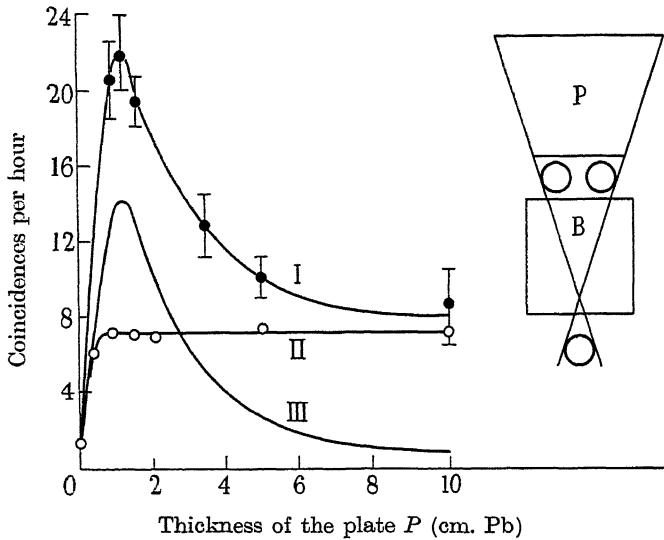


FIG. 2

lead block by the hard component would again cause coincidences. An unambiguous subtraction of the effect from the hard component seems therefore not to be possible at sea-level. By performing the experiments on high mountains or in the stratosphere, we would reduce the effect from the hard component since the intensity of the soft component increases very much relatively to that of the hard one with increasing height. On the other hand, the spectrum of the shower producing electrons would then not be the same. The spectrum above E_0 is fairly constant except for very high altitudes and E_0 would probably also be constant, but the value of α would presumably decrease since the hard component decreases relatively to the soft one. As will be seen from the following discussion, the exact

value of α has, however, no critical influence on most of the characteristic features of the theoretical curves. For the present we shall be content to compare our results with the experimental data now available, subtracting the hard component by extrapolation. We shall then find, as discussed in the following sections, that the agreement between theory and experiment is very satisfactory.

5. DISCUSSION OF THE ABSORPTION CURVE

We now discuss the absorption curve in lead and aluminium, that is the variation of the number of single particles, with the thickness of the material transversed. This is experimentally measured by placing the absorbing sheet between two of three counters placed in a vertical line. The curve

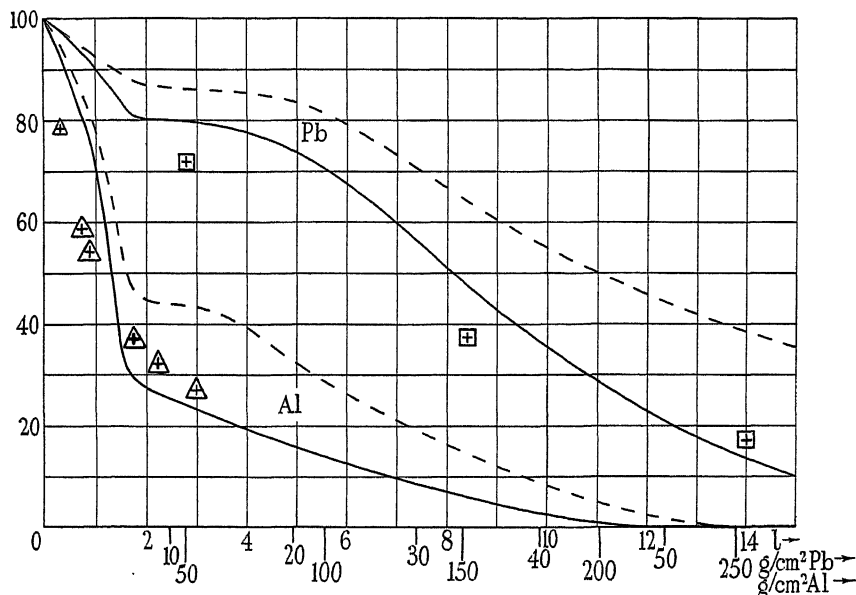


FIG. 3. Absorption curves for Pb and Al. ($l=1$ corresponds to 4.06 g./cm.^2 Pb and 18 g./cm.^2 Al.) Δ (Al) and \boxplus (Pb) are the experimental points of Auger, Leprince-Ringuet and Ehrenfest (1936). The full curves correspond to $\alpha = 10$ ($\sim \infty$), the dotted curves to $\alpha = 1$.

obtained represents the mean probability $\bar{P}(\geq 1) = \sum_{l=1}^{\infty} \bar{P}(N, l) = 1 - \bar{P}(0, l)$ for finding at least one particle below the plate if one particle falls on the top. In fig. 3 we give the theoretical curves for $\bar{P}(\geq 1)$ for Pb and Al as given

by (26) and Tables III and IV with $\alpha = 10$.* A variation of α from 10 to ∞ does not alter the curves appreciably, but as an illustration we have also given the corresponding curves for $\alpha = 1$ (dotted curves), and it is seen that the variation with α is so small that the absorption curves are practically independent of the spectrum of the soft component. Only for light elements might it be thought possible to determine α by very accurate absorption measurements. Owing, however, to the difficulties involved in subtracting the effect of the hard component in a proper manner (cf. § 4), we doubt whether this is practicable.

The beginning of the curves is determined by the absorption of the primary electron. For increasing thicknesses the secondary particles produced in the plate also contribute to $\bar{P}(\geq 1)$. For large thicknesses the curves are determined mainly by the absorption of the secondary particles. The remarkable form of the curves is, therefore, due to the combined effects of primary and secondary particles and not due to the special form of the spectrum used. If this form of the curve is not found experimentally, it only means that we cannot put all the geometrical factors equal to unity. In fact, by taking a larger value for the first geometrical factor than for the following ones we should obtain a smoother curve.

The curves obtained are roughly the same as those calculated by Heitler (1937*b*) with quite a different spectrum (given by (24) with $\alpha = 0$ and $\gamma = 1$ or 2) and neglecting the slow electrons. (This agreement is rather accidental and is due to the circumstance that the neglect of the slow electrons is more or less compensated by the increased importance of the greater energies in the spectrum.)

If plotted on a g./cm.^2 scale the two curves for Pb and Al are, to a first approximation, identical, so that the absorption follows roughly an ordinary mass absorption law as it was also found in the paper mentioned above. It is, however, a characteristic feature, which is also found in experiments, that the absorption coefficient is not constant, but much larger for the first 30–40 g./cm.^2 than for larger thicknesses. In fact the two values of the absorption coefficient, of about $15 \times 10^{-3} \text{ cm.}^2/\text{g.}$ and $5 \times 10^{-3} \text{ cm.}^2/\text{g.}$ for Al and $10\text{--}20 \times 10^{-3} \text{ cm.}^2/\text{g.}$ for the first 30–40 g./cm.^2 of Pb, are in very satisfactory agreement with the results of Auger, Leprince-Ringuet and Ehrenfest (1936) who find for the corresponding values 15×10^{-3} and $3.7 \times 10^{-3} \text{ cm.}^2/\text{g.}$ for Al and $13 \times 10^{-3} \text{ cm.}^2/\text{g.}$ for Pb.†

* Normalized so that $\sum_{N=0}^{\infty} \bar{P}(N, l) \equiv 100 \%$.

† This figure deviates from the one given by Auger and collaborators, 32 ± 3 for Pb. We have obtained our figures from fig. 2 of the authors.

In fig. 3 we have further given the two experimental curves found by Auger and collaborators which are also in agreement with those given by Maass (1936) and Sittkus (1938), the latter curves, however, lying somewhat lower. The general shape of the curves is seen also to be in very satisfactory agreement with the theoretical curves, taking into account the uncertainties of the experimental values due to the fact that the hard component is subtracted only by extrapolation.

The experiments of Gandin (1936) are, however, in contradiction both with those of Auger and collaborators and with our calculations. This discrepancy might be due to the fact that Auger's experiments are performed at Jungfraujoch (3500 m. above sea-level) and Gandin's at sea-level. We do not think, however, that this is so, because, as discussed in §4, the spectrum does not change considerably between sea-level and 4000 m. altitude, and apart from this the discrepancy between Gandin's results and our calculations still remains. Possibly the discrepancy may be explained, as pointed out by Jánosy,* by the presence of some absorbing material (roofs, etc.) above the apparatus in Gandin's experiments. In view of these discrepancies, new experiments on the absorption of the soft component seem therefore very desirable.

6. DISCUSSION OF THE ROSSI CURVE

Finally we discuss the Rossi (1933) coincidence curves, that is the variation with the thickness of the shower producing plate of the mean probability $\bar{P}(\geq N) = \sum_{N=0}^{\infty} \bar{P}(N, l)$ of finding at least N particles below the plate when one is falling on the top. N is given by the experimental arrangement and is usually 2 or 3. In fig. 4 we give the theoretical curves for $\bar{P}(\geq 2)$, $\bar{P}(\geq 3)$ and $\bar{P}(\geq 4)$ † for Pb as given by (26) and Tables III and IV with $\alpha = 10$. For illustration we have also plotted $\bar{P}(\geq 2)$ corresponding to $\alpha = 1$ (upper dotted curve) and $\alpha = \infty$ (lower dotted curve). In fig. 5 we give the corresponding curves for Al.

The first important feature which can be compared with experiments is the position of the maxima. For Pb this varies very little with N and lies at values of l between 3.5 and 4 (~ 14 – 18 g./cm.² Pb). For Al it lies for $\bar{P}(\geq 2)$ between 2 and 2.5 (~ 35 – 50 g./cm.² Al) and increases to between 3.5 and 4 (~ 55 – 75 g./cm.² Al) for $\bar{P}(\geq 3)$ and $\bar{P}(\geq 4)$. For other elements

* In a conversation with the author.

† Normalized so that $\sum_{N=0}^{\infty} \bar{P}(N, l) \equiv 100\%$.

we would obtain curves between the Pb and Al curves; e.g. for iron $\bar{P}(\geq 2)$ would have a maximum for l between 2.5 and 4 (~ 25 – 40 g./cm.² Fe). It should be noted that these data are essentially the same for all values of α between one and infinity so that we can conclude that the *position of the maxima is practically independent of the spectrum*. The corresponding experimental values are about 20 g./cm.² for $\bar{P}(\geq 2)$ in Pb (Schwegler 1935 and others), about 35 g./cm.² for $\bar{P}(\geq 2)$ in Fe (Pribsch 1935) and a very broad

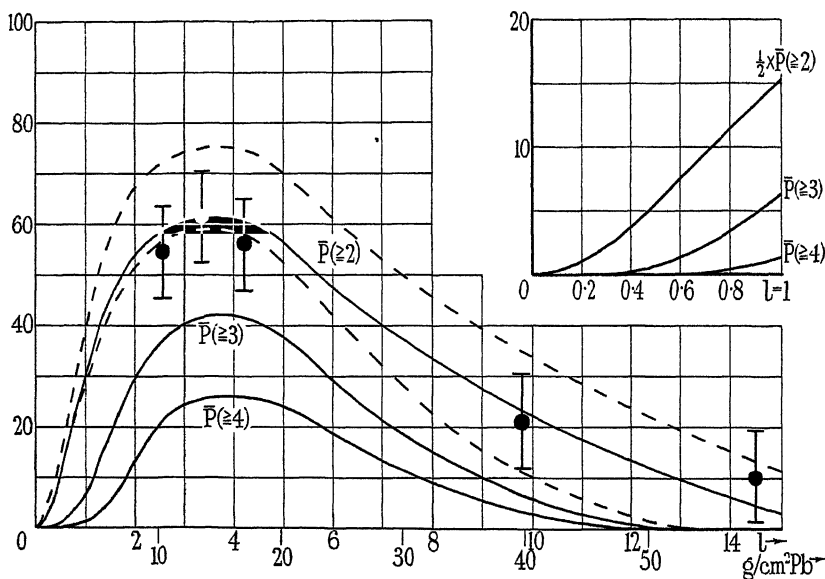


FIG. 4. Rossi curves for Pb ($l=1$ corresponds to 4.06 g./cm.² Pb). The experimental points for $\bar{P}(\geq 2)$ are those given by Schwegler (1935) (multiplied by a constant factor so as to make the maxima coincide). The upper dotted curve gives $\bar{P}(\geq 2)$ for $\alpha=1$, the lower one for $\alpha=\infty$. The figure in the corner gives the beginning of the Rossi curves on a larger scale.

maximum of $\bar{P}(\geq 3)$ in Al at 50 – 100 g./cm.² (Morgan and Nielsen 1936). (The shift of the maximum from $\bar{P}(\geq 2)$ to $\bar{P}(\geq 3)$ in Al seems not to have been observed, but we must remember that the Al curves are very flat and the shift, therefore, difficult to measure.) *The agreement between the theoretical and experimental values of the position of the maxima is very satisfactory.*

The next important feature to be compared with experiments is the relative magnitude of the maxima, on the one hand for different values of N but for the same element, and, on the other hand, for the same value of N but different elements. It seems that only few experiments are carried out

on these lines. Geiger and Zeiller (1937) have investigated $\bar{P}(\geq 2)$ and $\bar{P}(\geq 3)$ for Pb. Their values lead to the value 0.5 for $\frac{\bar{P}(\geq 3)_{\max}}{\bar{P}(\geq 2)_{\max}}$ in agreement with our value 0.7 for the same ratio. Morgan and Nielsen (1936) have investigated $\bar{P}(\geq 3)$ for Pb, Fe and Al. Their value is $\frac{\text{Pb}_{\max}}{\text{Al}_{\max}} \sim 10$. From figs. 4 and 5 we see that the theoretical $\frac{\text{Pb}_{\max}}{\text{Al}_{\max}} = 14$. This is in very good

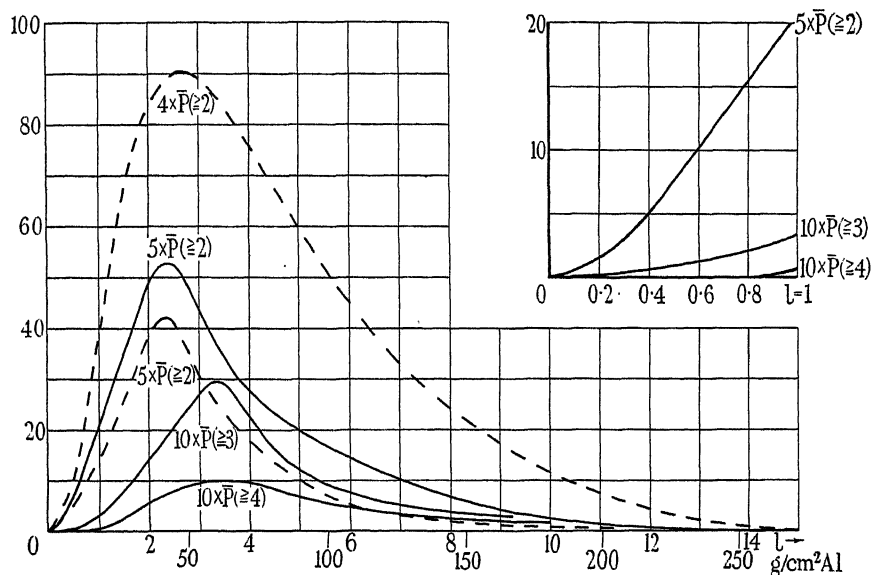


FIG. 5. Rossi curves for Al ($l=1$ corresponds to 18 g./cm.² Al). The upper dotted curve gives $\bar{P}(\geq 2)$ for $\alpha = 1$, the lower one $\bar{P}(\geq 2)$ for $\alpha = \infty$. The figure in the corner gives the beginning of the Rossi curves on a larger scale.

agreement with the experimental value. (For $\alpha = 1$ we would find 5.3 and for $\alpha = \infty$, 28. The accuracy of the measurements is, however, not high enough to allow one to determine α from this figure.)

In fig. 4 we have also plotted the experimental points for $\bar{P}(\geq 2)$ obtained by Schwegler (1935), using the arrangement mentioned in § 4. The curve has been multiplied with a constant factor so that the values at the maximum are identical. It is seen that the shape of the theoretical curve is in very good agreement with the experimental points.

If the mean number of soft particles crossing unit area per unit time were known at the place where the experiments are performed, it would also be possible to calculate the absolute number of coincidences per unit time

to be expected from a plate of given area. Of course the values obtained in this way would be rather difficult to compare with experiments, since the geometrical factors play a decisive role for the absolute number of coincidences (Zeiller 1935). If, however, our interpretation of the experiments of Clay and van Gemert is correct, we should expect the theoretical values for the absolute number of coincidences to be considerably smaller than the experimental ones due to the effect of the light quanta (cf. § 4).

The last feature which has been investigated in several experiments is the shape of the curves near the origin ($l = 1$) which is given separately in figs. 4 and 5. For $N = 2$ the curve increases practically linearly with the thickness, whereas for $N = 3$ the curve is more concave. This is in qualitative agreement with the experiments (Morgan and Nielsen 1935; Watase 1937). There is, however, a serious difficulty in the explanation of these experiments. It has been found that if the number of coincidences is plotted for different elements as a function of the thickness on a scale proportional to Z^2 (i.e. practically in l_0 units) the points all lie on the same curve (Hu Chien Shan 1937; Hu Chien Shan, Kisilbasch and Ketiladge 1937; Morgan and Nielsen 1937; Watase 1937). This means that the Rossi curves for small thicknesses should be independent of the critical energy of the shower producing materials, which is in contradiction to our results since the values for Al of both $\bar{P}(\geq 2)$ and $\bar{P}(\geq 3)$ are smaller by roughly a factor 10 than the corresponding values for the same l in Pb. This can easily be understood in cases where only two particles are necessary to produce a coincidence. In a thin sheet each primary light quantum produces a pair and what is measured is then simply the pair production which is known to be proportional to Z^2 . In the experiments mentioned above coincidences of at least three particles were measured. The proportionality with Z^2 can in this case only be explained by a simultaneous arrival of a light quantum and an electron. As discussed above, the probability for this happening does not seem to be at all negligible (§ 4).

The moment, therefore, that a thin sheet of material is placed above the counters the light quantum will produce a pair which together with the electron produces a coincidence. As the probability of this process is proportional to Z^2 and to the thickness (as long as this is small), we should expect just the curve obtained experimentally. For larger thicknesses the ordinary showers become more important and we should then expect a deviation both from the Z^2 law and from the linearity. In fact the curves are found to be concave upwards in agreement with our interpretation, but the deviation from the Z^2 law seems not to have been found for thicknesses up to $l = 1$. Furthermore, we have to remember that for small

thicknesses the effect of the hard component is not known at present. The experiments of Schwegler (1935) mentioned in § 4 seem to show that the number of coincidences due to the hard component increases very steeply at the beginning of the curve (cf. fig. 2). It is, therefore, quite possible that for small thicknesses the hard component cannot be neglected. It would be very important to know whether the hard component produces showers at *small* thicknesses and how they depend upon the thickness and Z . To clarify this point, further experiments are necessary.

In conclusion I wish to express my appreciation to Dr W. Heitler for suggesting the problem to me and my thanks for many stimulating and helpful discussions in the course of the calculations. I also wish to thank Professor P. M. S. Blackett and his collaborators for valuable discussions on many of the experimental points. Furthermore, I should like to thank Professor N. Bohr for kindly permitting me to be absent on leave from the Copenhagen Institute of Theoretical Physics, and Professor N. F. Mott for giving me the opportunity to work in the University of Bristol.

APPENDIX I

In order to calculate the functions $f(l, y)$ given by

$$\left. \begin{aligned} f(l, y) &= \sum_{n=1}^{\infty} f_n(l, y), \\ f_n(l, y) &= \frac{(2 \cdot 0 \cdot 6 \cdot \log 2)^n}{2(n!)^2(n-1)!} \int_0^l dl' \{e^{-0 \cdot 6(l-l')} l'^n (l-l')^{n-1}\} \left[\frac{\int_0^{y'} dt e^{-t} (y-t)^n t^{l'+n-1}}{(l'+n-1)!} \right], \end{aligned} \right\} \quad (28)$$

for small values of l it is possible to expand every $f_n(l, y)$ in powers of l . (It is easily seen that the series begins with l^{2n} .) As the coefficients are cumbersome to work out and involve complicated integrals, we have proceeded in the following way: Using the formula for the [] in (28)

$$\left. \begin{aligned} [n, y, l'] &= \sum_{s=0}^n \binom{n}{s} (-1)^s y^{n-s} (l' + n)_s W(l' + n + s, y), \\ (n)_s &= n(n+1) \dots (n+s-1), \end{aligned} \right\} \quad (29)$$

we can easily compute $[n, y, l']$ as functions of l' for small values of n and the different values of y by means of Pearson's Tables of the incomplete

Γ -function.* In the region $0 \leq l' \leq 2$ all the $[n, y, l']$ curves can be approximated by one or two straight lines the coefficients of which can be measured on the graphs, and the integrations over l' can then be performed analytically. The results are given in Table I. (The values for $l > 2$ are obtained by interpolation between our values and those calculated by Miss Eriksen for $l \geq 10$ (Table I), the interpolations being most conveniently performed on graphs giving $\log f(l, y)$ as functions of $\log l$ for constant values of y or $\log y$ for constant values of l .)

APPENDIX II

The number of electrons which at l have energies between $\frac{E}{E_0} \left(> \frac{E_c}{E_0} \right)$ and $\frac{E}{E_0} + \Delta \frac{E}{E_0}$ is given by

$$E_0 p(l, E) \Delta \frac{E}{E_0} = - \frac{\partial f(l, y)}{\partial E} \Delta E = e^y \frac{\partial f(l, y)}{\partial y} \Delta \frac{E}{E_0}, \quad y = \log \frac{E_0}{E}. \quad (30)$$

TABLE V. THE ENERGY SPECTRUM $E_0 p(l, E)$ FOR THE SECONDARY ELECTRONS AS A FUNCTION OF l AND $y = \log E_0/E$

$l \backslash y$	2	4	6	8	10	15
0.2	0.0467	0.423	3.25	23.6	179	2.82×10^4
0.4	0.177	1.63	13.1	96.0	715	1.16×10^5
0.6	0.363	3.64	29.1	224	1720	2.86×10^5
0.8	0.591	6.31	53.1	427	3420	6.01×10^5
1	0.844	9.87	85.8	718	5880	1.08×10^6
2	2.16	36.6	420	4230	4.13×10^4	1.08×10^7
3	3.06	74.3	1050	1.508×10^4	1.75×10^5	6.23×10^7
5	2.32	131	4070	7.15×10^4	1.02×10^6	8.77×10^8
10	0.224	66.5	6110	3.06×10^5	9.74×10^6	2.41×10^{10}

We have calculated $E_0 p(l, E)$ from the values of $f(l, y)$ obtained in Appendix I, by using that (30) can also be written in the form

$$E_0 p(l, E) = \frac{e^y}{y} f(l, y) \frac{\partial \log f(l, y)}{\partial \log y}, \quad (31)$$

where the differential coefficients can easily be measured on the graphs giving $\log f(l, y)$ as functions of $\log y$ for constant values of l , since in this

* I should like to thank Miss Bodil Eriksen of the Copenhagen Institute for Theoretical Physics for much help in these numerical calculations. Miss Eriksen has also calculated by direct numerical integrations the values of $f(l, y)$ for $l \geq 10$ given in Table I.

picture the curves are roughly parallel straight lines. The values obtained for the differential coefficients in (31) vary, therefore, only slowly with both l and y so that the form of the curves for $E_0 p(l, E)$ is mainly determined by the factor $e^y f(l, y)/y$ and will have maxima in respect to l for nearly the same l -values as $f(l, y)$. In Table V we give the results.

SUMMARY

We have made further calculations on the cascade theory of showers put forward by Bhabha and Heitler and by Carlson and Oppenheimer. In §1 the number of "slow" electrons in a shower is calculated, i.e. of electrons having an energy less than the critical energy of the shower producing material. In §2 we discuss the fluctuations of the number of particles about the mean number. By means of a simplified model for the multiplication process it is made plausible that the Poisson formula holds for the fluctuation in small showers. In §3 the energy spectrum of the soft component is discussed. With this energy spectrum the mean probabilities of finding one, two or more particles emerging from a certain layer are calculated, giving the absorption curve and the Rossi transition curves. In §§4-6 we compare our results with experiments. *The quantitative agreement between theory and experiment for the absorption and the Rossi curves is very satisfactory* both with regard to the shape of the curves and their dependence on the material. There is, however, a difficulty in understanding the Z^2 law found for the Rossi curve at small thicknesses. This is possibly explained by the simultaneous effects of a light quantum and an electron. The spread of showers from air is discussed in this connexion.

REFERENCES

- Auger, Leprince-Ringuet and Ehrenfest 1936 *J. Phys. Radium*, 7, 58.
 Auger and Maze 1938 *C.R. Acad. Sci., Paris*, 207, 228.
 Bethe and Heitler 1934 *Proc. Roy. Soc. A*, 146, 83.
 Bhabha and Heitler 1937 *Proc. Roy. Soc. A*, 159, 432.
 Blackett 1938 *Proc. Roy. Soc. A*, 165, 11.
 Carlson and Oppenheimer 1937 *Phys. Rev.* 51, 220.
 Clay and van Gemert 1936 *Physica*, 3, 763.
 Euler 1938 *Z. Phys.* (in the press).
 Furry 1937 *Phys. Rev.* 52, 569.
 Gandin 1936 *Ric. Sci.* VII, II, Nos. 3-4.
 Geiger 1937 *Phys. Z.* 38, 936.
 Geiger and Zeiller 1937 *Z. Phys.* 105, 517.
 Heitler 1936 "Quantum Theory of Radiation." Oxford.
 — 1937a *Proc. Roy. Soc. A*, 161, 261.

- Heitler 1937*b* *Nature, Lond.*, **140**, 235.
 — 1938 *Proc. Roy. Soc. A*, **166**, 529.
 Hilgert and Bothe 1936 *Z. Phys.* **99**, 353.
 Hu Chien Shan 1937 *Proc. Roy. Soc. A*, **158**, 581.
 Hu Chien Shan, Kisilbasch and Ketiladze 1937 *Proc. Roy. Soc. A*, **161**, 95.
 Maass 1936 *Ann. Phys., Lpz.*, **27**, 507.
 Morgan and Nielsen 1935 *Phys. Rev.* **48**, 773.
 — — 1936 *Phys. Rev.* **50**, 882.
 — — 1937 *Phys. Rev.* **52**, 568.
 Priebsch 1935 *Z. Phys.* **95**, 102.
 Rossi 1933 *Z. Phys.* **82**, 151.
 Schmeiser and Bothe 1938 *Ann. Phys., Lpz.*, **32**, 161.
 Schwegler 1935 *Z. Phys.* **96**, 62.
 Sittkus 1938 *Z. Phys.* **108**, 421.
 Watase 1937 *Proc. Phys. Math. Soc. Japan*, **19**, 749.
 Zeiller 1935 *Z. Phys.* **96**, 121.
-

The specific heat of β -brass

BY R. EISENSCHITZ

(Communicated by Sir William Bragg, O.M., P.R.S.—Received 26 July 1938)

1. INTRODUCTION

The specific heat of β -brass as measured by Sykes (1935) and by Moser (1936) shows an anomalous rise beginning at 160° C, the slope increasing markedly with increasing temperature. At 463° C a sharp maximum—a λ -point—is reached and the specific heat drops abruptly to low values, the whole drop taking place within 7°. Moser states that this drop is steep but continuous. Both authors obtain a slope of 0.03 cal./degree² g. approximately by interpolating between their experimental values. The thermal expansion coefficient goes parallel with the specific heat and has a maximum of similar shape (Steinwehr and Schulze 1934).

This phenomenon is explained as the transition of the Cu and Zn atoms from an ordered arrangement to disorder.

The existing theories account for the characteristic λ -shape of the specific heat curve, but the quantitative agreement between theory and experiment is rather poor. The maximum specific heat and the slope of the specific heat curve below the maximum, as given by the theories, are too small.

It is conceivable that these discrepancies might disappear if better approximations were used in the existing theories. An attempt is made in the present paper to obtain better agreement by introducing into the theory the effect due to the thermal expansion which has not been taken into account so far.

A recent investigation (Müller 1938) of the rotation of molecules in solids has shown this phenomenon to be sensitive to change of volume. It seems therefore not unlikely that the neglected volume change may be responsible for the disagreement between experiment and theory in the case of β -brass.

The drop in the specific heat beyond the λ -point is discontinuous according to the existing theories, which is not in agreement with the experiments. It is shown in the present paper that this discontinuity can be removed by a modification of the fundamental assumptions to be discussed in the following.

2. REVIEW ON THE EXISTING THEORIES

The theory of order-disorder transition for an equimolecular binary alloy in a cubic lattice was developed on three different lines by Bragg and Williams (1934), by Bethe (1935) and by Kirkwood (1938). In the three theories the specific heat is found to be discontinuous at a "critical temperature"; volume changes are neglected. In the following the origin of the discontinuity is discussed. A complete report on the order-disorder transition is given by Nix and Shockley (1938).

The original theory of Bragg and Williams is based on the assumption that the ordering energy is proportional to the long-distance order of the lattice. If the atoms are assumed to interact within small ranges only, this assumption seems to be somewhat arbitrary. Indeed the authors claim that the theory gives only a general idea of the way in which the long-distance order varies with temperature. The commencement of order, as they point out, depends in a very sensitive way on the relation between order and ordering energy. For these reasons it may be concluded that the theory of Bragg and Williams is not sufficient to decide whether the specific heat is discontinuous or not.

The theory of Bethe starts from a well-defined molecular model. A lattice is assumed to consist of two kinds of atoms (A and B). The total energy is the sum of a potential energy arising from the arrangement of the atoms and the vibrational energy which is assumed to be independent. The theory does not deal with the vibrational energy. The energy of arrangement is assumed as the sum of interaction energies of pairs of nearest neighbours

$\phi_{AA}, \phi_{BB}, \phi_{AB}$; $\phi = (\frac{1}{2})(\phi_{AA} + \phi_{BB}) - \phi_{AB} > 0$. This makes the energy E_m of any configuration with m links between atoms of the same kind

$$E_m = m\phi. \quad (1)$$

All possible arrangements have the same statistical weight.

These assumptions determine all thermodynamical properties such as the specific heat. To calculate them is only a mathematical problem involving the enumeration of all configurations with a given number of links between atoms of the same kind.

Bethe gives the rigorous solution for the one-dimensional chain of n atoms. The partition function is found to be

$$P = (1+x)^n, \quad x = e^{-\phi\tau}, \quad \tau = 1/kT.$$

This partition function corresponds to a chain in which the numbers of A and B atoms vary between 0 and n . It will be shown in § 5 that the same partition function holds for a chain with $n/2$ A atoms and $n/2$ B atoms.—For a “central atom” in a two- or three-dimensional lattice, which interacts with its z nearest neighbours (“first shell”) only, the partition function

$$P = (1+x)^z$$

is obtained, and a similar partition function is found, which takes account also of the interaction of the atoms of the first shell with their nearest neighbours (“second shell”). The specific heat curves have broad maxima and no discontinuity.

The partition function for the complete two- or three-dimensional lattice is extremely difficult to calculate. In section 6 of his paper Bethe works out a method of approximation. He divides the lattice points in two groups (a and b positions) which form two interpenetrating lattices. He defines as “ R ” (right) atoms A atoms in an a position or B atoms in a b position, as “ W ” (wrong) atoms A atoms in a b position or B atoms in an a position. A central atom and the various shells round it are considered, and it is assumed (in first approximation) that the only effect of the outer shells is to make R atoms in the first shell more likely than W atoms; this follows from the assumption that the outer shells themselves contain more R than W atoms. In order to calculate the partition function according to these assumptions, a mean energy for each W atom in the first shell is introduced, which accounts for the interaction of this atom with the remaining atoms in the infinite lattice. This energy is put $-(1/\tau)\log \epsilon$, and is assumed to depend upon the temperature. The energy of interaction for every R atom in the first shell is assumed to be 0. It is shown that the relative probabilities

r_n and w_n , that n atoms of the first shell are W and the central atom is R or W respectively, are

$$r_n = \binom{z}{n} x^n \epsilon^n, \quad w_n = \binom{z}{n} x^{z-n} \epsilon^n.$$

From this the probabilities are calculated that the central atom is W or that any atom of the first shell is W . These probabilities must be equal, for the central atom is in no way distinguished from the others. By equating them ϵ is computed as a function of the temperature. Two different analytical expressions are obtained for ϵ , one holding below, the other beyond the critical temperature; this results in a discontinuity of the specific heat.

Bethe's method is subject to criticism. If according to the original assumptions the energy of every configuration depends only upon the number of links between atoms of the same kind and all configurations have the same statistical weight, no difference can be made between atoms in a or b positions. The probability that any lattice point is occupied by an R or a W atom must be equal at all temperatures. The interaction energy of an atom in the first shell with the rest of the infinite lattice is the same for an R or W atom, since one type is transformed into the other by interchanging the a and b positions. There are two types of long-distance order, with the majority of A atoms in a or b positions respectively. Both types have to occur in thermodynamical equilibrium with equal probability. In Bethe's treatment "right" and "wrong" atoms are defined relative to the first type of order, and his assumption that the outer shells contain more R than W atoms implies that transitions to the other type of order should not occur. At temperatures sufficiently below the critical such transitions are improbable and Bethe's treatment may be an adequate approximation. Near the critical temperature where all long distance order is breaking down, the two types of order must be expected to occur frequently. Bethe's method is not justified near the critical temperature and cannot prove the discontinuity of specific heat.

Kirkwood starts from the same molecular model as Bethe. He classifies the arrangements of atoms according to their degree of long-distance order, as defined by Bragg and Williams. To every class of configurations an energy distribution is attached, which expresses the number of configurations for given long-distance order and given energy. According to mathematical statistics every distribution function of energy is uniquely determined by the totality of the moments of energy. Kirkwood calculates the moments of the order 0 to 3 as functions of the degree of order and shows that all distribution functions have steep maxima. The distributions may therefore

be roughly approximated assuming all configurations of one class to have the mean energy of this class. In this way a first approximation to the partition function is obtained. The specific heat is calculated by substituting, for the sum in the partition function its maximum term, an approximation which is legitimate in the limit of an infinite number of atoms. The result is identical with the result of Bragg and Williams. Their assumptions, arbitrary at first sight, are shown to be the correct first approximation to a kinetic theory. A second approximation is obtained by assuming the distribution functions to be error functions. The result is somewhat similar to Bethe's but not identical with it.

The discontinuity of specific heat is derived in Kirkwood's theory by substitution of the maximum term for the sum in the partition function, in other words in the—experimentally not realizable—limit of an infinite lattice. Disregarding this limitation it may therefore appear as if the discontinuity were derived from a molecular model without any arbitrary hypothesis. Kirkwood's theory, however, is only an approximation. It introduces an apparent long-distance interaction which makes the drop of specific heat more abrupt than it really is. Applying this method to the one-dimensional chain, it can be shown to have a discontinuity of specific heat at the critical temperature $T = k/\phi$, which is in contradiction to Bethe's rigorous calculation.

Kirkwood's theory seems nevertheless to be a good approximation if it is applied to a finite number of atoms. In this way the discontinuity is avoided. The introduction of a finite number of atoms is not more arbitrary than that of an infinite number.

In order to account for the additional thermal expansion it is necessary to make the interaction energies functions of the volume. This involves the introduction of at least two parameters.

In the following the specific heat curve is derived by applying Kirkwood's first approximation, essentially the theory of Bragg and Williams, to a finite lattice with the interaction energies depending on the volume.

3. THE ORDER-DISORDER TRANSITION AT CONSTANT PRESSURE

The theory of Bragg and Williams is first to be modified in such a way as to hold for a finite lattice of j atoms at constant volume.

The lattice points are classified into a and b points forming two interpenetrating simple lattices. In any configuration let the number of a points

occupied by A atoms be p . (The degree of order as defined by Bragg and Williams is $S = -1 + 4p/j$.) The partition function

$$P = \sum_{p=0}^{\frac{1}{2}j} e^{-6j\phi\tau(1-2p/j)(2p/j)} \left(\frac{\frac{1}{2}j}{p}\right)^2, \quad (\tau = 1/kT), \quad (2a)$$

is transformed by applying Stirling's formula, omitting constant factors and introducing the variable $x = -\frac{1}{2} + 2p/j$ into

$$P = \sum_{x=0}^{\frac{1}{2}} \exp \left[6j\phi\tau(x^2 - \frac{1}{4}) - (\frac{1}{2}j) \{ (1+2x) \log(1+2x) + (1-2x) \log(1-2x) \} \right], \quad (2b)$$

in which sum x assumes the values $0, 2/j, 4/j, \dots, 1/2$. Substitution of the maximum term for the partition function would give the formulae of Bragg and Williams. In order to evaluate the sum at the critical temperature, the lattice being finite, the exponent is expanded in powers of x and the summation is approximated by integration over the variable $y = x^2$ between the limits 0 and ∞ :

$$\begin{aligned} P &= \int_0^\infty \exp \left[j \left\{ -(\frac{3}{2})\phi\tau + y(6\phi\tau - 2) - y^2(\frac{4}{3}) \right\} \right] dy \\ &= \exp \left[j \left\{ -(\frac{3}{2})\phi\tau + (\frac{3}{16})(2 - 6\phi\tau) \right\} \right] [1 + \Theta \{ \frac{1}{8}(3j)^\frac{1}{2} (6\phi\tau - 2) \}] \end{aligned} \quad (2c)$$

with the error integral $\Theta(t) = (2/\sqrt{\pi}) \int_0^t e^{-s^2} ds$.

At the critical temperature $\phi\tau = \frac{1}{3}$. Putting $\vartheta = \phi\tau - \frac{1}{3}$,

$$P = 1 + \alpha\vartheta + \beta\vartheta^2 + \gamma\vartheta^3 + \dots \quad (2d)$$

is obtained, with

$$\alpha = 3(3j/\pi)^\frac{1}{2}, \quad \beta = \frac{27}{4}j, \quad \gamma = 27(3j^3/4\pi)^\frac{1}{2}, \quad (3)$$

and the mean energy and its derivatives at the critical temperature ($\vartheta = 0$) are

$$\bar{E} = \frac{1}{2}(3j\phi) - \alpha\phi, \quad (4a)$$

$$d\bar{E}/dT = k(2\beta - \alpha^2)/9, \quad (4b)$$

$$d^2\bar{E}/dT^2 = -(k/T) \left[\frac{2}{3}(2\beta - \alpha^2) + \frac{1}{27}(6\gamma + 6\alpha\beta - 2\alpha^3) \right]. \quad (4c)$$

The specific heat per gramme atom, C_v , and its derivative are

$$C_v = 0.590 Lk,$$

$$dC_v/dT = -(Lk/T)(1.18 + j^\frac{1}{2}0.40)$$

accordingly, where L denotes Loschmidt's number. In the $\lim j \rightarrow \infty$ the slope of the specific heat curve tends to infinity with $j^\frac{1}{2}$.

The above theory is to be modified so as to account for the specific heat of thermal expansion. This is done by making the interaction energies depend upon the volume and to take an average of the partition function over all volumes. The theory is simplified by assuming the vibrational energy to be independent of the energy of arrangement. This eliminates the effect of fluctuations of density; such configurations are only to be taken into account in which all elementary cubes have the same volume.

Regarding the relation of atomic energies and the volume, it is known that the volume at low temperatures must make ϕ_{AB} a minimum and that by expanding the lattice tends towards the minimum of $(\frac{1}{2})(\phi_{AA} + \phi_{BB})$. The simplest mathematical formulation for potential energy curves near their minima is made by assuming ϕ_{AB} and $(\frac{1}{2})(\phi_{AA} + \phi_{BB})$ to depend upon the square of the volume change.

If u' is the volume and u'' and u''' the volumes, at which $(\frac{1}{2})(\phi_{AA} + \phi_{BB})$ and ϕ_{AB} have their respective minima, $u'' > u'''$, and

$$u = (u' - u''')/(u'' - u'''),$$

the interaction energies are assumed to depend on the volume in the following way:

$$(\frac{1}{2})(\phi_{AA} + \phi_{BB}) = \chi(u) = \phi[(1-a) + a(1-u)^2], \quad (5a)$$

$$\phi_{AB} = \psi(u) = \phi bu^2, \quad (5b)$$

a and b being constants. For $a=b=0$ Bethe's model is obtained.

In configurations with pA atoms occupying a points the energy is, according to the approximation of Bragg and Williams,

$$E_p = 3j[2(2p/j)(1-2p/j)\chi + \{4p^2/j^2 + (1-2p/j)^2\}\psi]$$

or

$$E_x = 6j[\frac{1}{4}(\psi + \chi) + x^2(\psi - \chi)],$$

and the partition function

$$P = 2 \int du \sum_{x=0}^{\frac{1}{2}} \exp [j\{-(\frac{3}{2})(\chi + \psi)\tau + 6x^2(\chi - \psi)\tau - (\frac{1}{2} + x) \log(1+2x) - (\frac{1}{2} - x) \log(1-2x)\}].$$

The limits of the integral are approximated by $-\infty$ and ∞ .

If χ and ψ are introduced as functions of u according to equ. (5a) and (5b) and the integration is carried out, the partition function is found to be

$$P = \sum_{x=0}^{\frac{1}{2}} \left(\frac{16\pi}{6j\phi\tau[(a+b) - 4x^2(a-b)]} \right)^{\frac{1}{2}} \exp \left[\frac{6j\phi\tau a^2(\frac{1}{4} - x^2)^2}{\frac{1}{4}(a+b) - x^2(a-b)} + 6j\phi\tau(x^2 - \frac{1}{4}) - (\frac{1}{2}j) \{(1+2x) \log(1+2x) + (1-2x) \log(1-2x)\} \right].$$

The first factor may be omitted as being approximately constant. The partition function is evaluated for temperatures below the critical by substitution of its maximum term for the sum, in the neighbourhood of the critical temperature by expansion and integration in analogy to equ. (2c).

It is convenient to put

$$a^2/(a+b) = r, \quad (a-b)/(a+b) = s, \quad 1-r(2-s) = A,$$

$$8r(1-s)^2 = B, \quad 48rs(1-s)^2 = C, \quad \frac{4}{3} - B/A = G,$$

$$\log [(1+2x)/(1-2x)] = \xi, \quad (1-4x^2)/(1-4sx^2) = \eta;$$

$$1-2r\eta + rs\eta^2 = \zeta, \quad 2\zeta + 32rx^2(1-s\eta)^2/(1-4sx^2) = \omega.$$

The result of the maximum method is

$$\phi\tau = \frac{1}{12}\xi/x\zeta, \quad (6a)$$

$$C_p = \frac{8Lk(6\phi\tau)^2 x^3 \eta^3}{[8x\zeta/(1-4x^2) - \xi\omega]}. \quad (6b)$$

Near the critical temperature, in the $\lim x = 0$ accordingly, these equations become

$$\phi\tau = \left(\frac{1}{3}A\right) \left[1 + Gx^2 + \left\{\frac{16}{5} - \frac{4B}{3A} + \left(\frac{B^2}{A^2} - \frac{C}{A}\right)\right\}x^4\right], \quad (7a)$$

$$C_p = Lk\left(\frac{2}{G}\right) \left[1 - \left\{\frac{16}{5G} - \frac{4B}{3AG} + \frac{B^2}{A^2G} - \frac{C}{AG}\right\}x^2\right]. \quad (7b)$$

The critical temperature is given by $x = 0$, $\phi\tau = 1/3A$, and the specific heat is $C_p(T_0) = 2k/G$.

Evaluation of the partition function by expansion and integration gives the specific heat and its gradient at the critical temperature

$$C_p(T_0) = 2Lk \left[\frac{1-2\pi}{G} + \frac{B}{AG^{\frac{1}{2}}(j\pi)^{\frac{1}{2}}} \right], \quad (8)$$

$$\frac{dC_p}{dT_0} = -\frac{2C_p}{T} - \frac{Lk}{T} \left[\frac{9B^2}{2G^{\frac{3}{2}}A^2(j\pi)^{\frac{1}{2}}} + (1-2/\pi) \frac{6B}{AG^2} + \frac{4j^{\frac{1}{2}}(4/\pi-1)}{G^{\frac{3}{2}}} \right]. \quad (9)$$

In order to construct the specific heat curve, the specific heat according to equ. (6) and a straight line satisfying equ. (8), and (9) are plotted; the first is assumed to be valid below, the second above, their intersection. The arithmetic mean of the specific heats of Cu and Zn as given by Moser is added to account approximately for the specific heat of vibration.

The curve (1) in fig. 1 is constructed with the following constants: $T_0 = 740^\circ \text{K}$, $a = 0.225$, $b = 0.203$, $j = 10^4$, and the specific heat of vibration as given by curve (2). From these data $dC_p/dT_0 = 1.66 \text{ cal./degree}^2 \text{ g. atom}$ and $\phi = 4.38 \cdot 10^{-14} \text{ erg}$ are calculated.

In fig. 1 the experimental values of Sykes and Moser are plotted as crosses and circles respectively; the specific heat per gramme given by these authors is multiplied by 64 as mean atomic weight. For the sake of comparison, Kirkwood's second approximation is plotted in the same figure (curve 3). The latter curve shows poor agreement only with experimental data, curve (1) agrees fairly well.

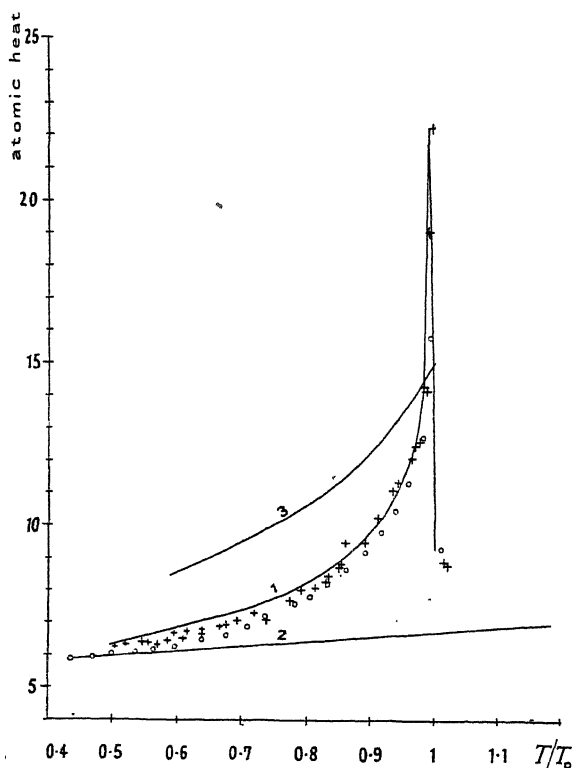


Fig. 1. 1, Theoretical C_p curve. 2, specific heat of vibrations. 3, C_p curve according to Kirkwood. +, measurements by Sykes. O, measurements by Moser.

It is seen that the parameters can be adapted so as to cover the experimental facts.

It may be noticed that two significant experimental quantities, the maximum specific heat and the slope of the specific heat curve below the

maximum, are determined in the theory by fixing the one constant G only. As a and b have the same magnitude, G is mainly determined by their arithmetic mean and is not very sensitive to changes of their differences. The same holds for the calculated specific heat curve. The theory can be tested, therefore, if a or b is derived from experimental data. b as curvature of the Cu-Zn energy curve can be related to the compressibility of β -brass at room temperature.

The relative expansion from room temperature to the critical temperature is according to Steinwehr and Schulze (1934) $\Delta v = 0.033$; the compressibility is $\kappa = 10^{-6}$ atm. $^{-1}$; the volume of a gram atom is $V = 8.1$ cm. 3 . For the elastic energy of expansion the equation holds:

$$\left(\frac{1}{2}\right)(\Delta v)^2 V / \kappa = 3L\phi bu^2.$$

At the critical temperature $u \cong \frac{1}{2}$ and with the computed value of ϕ , $b = 0.29$ is found instead of 0.203. This gives a check for the magnitude of b .

The validity of the present theory is limited by the assumption of independent energies of vibration and atomic arrangement. It cannot account for the fact that the normal lattice expansion is larger than the expansion due to the transition.

The present calculation shows that even a small expansion may give an appreciable difference between C_p and C_v .

The parameter j enters in the decreasing branch of the specific heat curve only. If its numerical value were applied to the calculation of C_v with $T_0 = 740^\circ$ K, $dC_v/dT_0 = -0.11$ cal./degree 2 g. atom would result, a slope of about 1/20 of dC_p/dT_0 . C_v would accordingly not even approximate to a discontinuity. Experimental data on C_v are lacking. An attempt is made in the following sections to obtain information about C_v from the molecular model.

4. A GENERAL THEOREM OF STATISTICAL MECHANICS

The kinetic theory of discontinuities of thermodynamical quantities is so far not satisfactory. The existing theories of phase transformations use molecular models in order to calculate the free energies of the virtually separated phases, but the discontinuity is introduced by application of the thermodynamical condition of equilibrium and not derived from the mechanical equations. The existing theories on discontinuities of specific heat are making use of assumptions for which there is no *a priori* evidence. This was shown in § 2 with regard to the order disorder transition. For the

ferromagnetic Curie point a discontinuity was derived by the theories of Weiss and of Heisenberg. These theories follow similar lines as the theories of Bragg and Williams, and Kirkwood respectively, and are subject to similar criticism. Recent contributions to the problem of sharpness of the Curie point were made by Opechowsky (1937) and Mott and Potter (1937). Only Van der Waals's theory of evaporation and recent theories on the same subject (for instance by Lennard-Jones and Devonshire 1937) attempt to derive discontinuities from molecular models without any additional hypothesis.

The discontinuities cannot be explained by the special kind of forces acting in the system. For substances with different kinds of molecular forces have the same discontinuities (melting for instance) and one and the same substance may have a continuous or discontinuous transition according to external conditions (evaporation above and below the critical pressure). The Hamiltonian influences the thermodynamical properties in an indirect way only, as shown in the fundamental theorems of statistical mechanics.

We shall discuss the statistical mechanics of a system of molecules in a box with the volume V . The possible energy states of the system E_n (with the weight w_n) are depending on the volume.

The partition function

$$P(\tau, V) = \sum_n e^{-E_n(V)\tau} w_n$$

(where $\tau = 1/kT$) is related to the thermodynamic quantities F = free energy and \bar{E} = mean energy by

$$F = -\log P/\tau, \quad \bar{E} = -\partial \log P / \partial \tau. \quad (10)$$

The coefficients w_n being statistical weights are essentially positive.

In the classical limit the partition function is

$$P(\tau, V) = \int e^{-E\tau} w(E, V) dE,$$

with the weight function $w(E, V) \geq 0$. In this approximation the partition function is factorized, one factor depending on the kinetic and one on the potential energy only, of which factors the first is a continuous function of temperature, independent of the volume and cannot possibly lead to any discontinuity.

If the molecules are assumed to be enclosed in a cylinder with a piston

carrying such a weight that the pressure is p , we can formulate the statistical mechanics of systems at constant pressure. The partition function is

$$P(\tau, p) = \int dV \sum_n e^{-[E_n(p) + pV]\tau} w_n,$$

or

$$P(\tau, p) = \int dV \int e^{-(E + pV)\tau} w(E, p) dE,$$

respectively. The free energy at constant pressure and the enthalpy is obtained by introducing P into formulae (10). The only difference is that here the parameter is p instead of V . The integration over V is to be carried out to a large but finite limit only, in order to avoid infinity of co-ordinates, but it follows from all known thermodynamical facts that the partition function becomes independent of this limit if the latter tends to infinity.

From these equations it is seen that the Hamiltonian enters only indirectly to the thermodynamical quantities through the weight function. Although every Hamiltonian defines its weight function uniquely, an infinite multiplicity of virtual Hamiltonians may give the same weight function.*

All systems with discontinuities may accordingly have very different kinds of forces but will probably have weight functions of characteristic properties. In trying to find them a general theorem could be established:

(I) For any physically possible ergodic system with all parameters fixed and in thermal equilibrium, the mean energy and all its derivatives are continuous functions of the temperature.

Proof. According to the assumptions we have

$$P(\tau) = \sum_n e^{-E_n \tau} w_n = e^{-E_1 \tau} \sum_n e^{-(E_n - E_1) \tau} w_n$$

$$(w_n > 0), \quad (E_n - E_1 \geq 0).$$

This sum converges for finite temperature. We assume τ to be a complex variable $\tau = \sigma + i\rho$. The sum converges accordingly for $\rho = 0$, $\sigma \geq \alpha > 0$. As $w_n > 0$ the convergence is absolute. If $\rho \neq 0$ the sum converges as well, because every term of the absolutely convergent sum is multiplied with a number the absolute value of which is one. As

$$|e^{-(E_n - E_1) \tau}| \leq e^{-(E_n - E_1) \alpha},$$

* For any empirical equation of state it is accordingly possible to construct many mechanical models. With an empirical T - \bar{E} curve at constant volume, it is even possible to construct a one-dimensional model.

the convergence is uniform relative to τ on any closed contour within the range

$$0 < \alpha \leq \sigma \leq \beta, \quad -\gamma \leq \rho \leq \delta,$$

where β, γ, δ are positive numbers. As all terms of the sum are analytic functions of τ in the interior of the contour, the sum in the partition function is an analytic function of τ ; for series converging uniformly on a closed contour, the terms of which are analytic functions within the contour, are also analytic functions. The partition function, a product of this sum and an exponential, is analytic as well. The function \bar{E} is a fraction of analytic functions. As the denominator $P(\tau) > 0$ it is analytic; \bar{E} and its derivatives are accordingly continuous on the real axis.

This theorem excludes discontinuities of the mean energy or derivatives in any system in equilibrium, in which thermodynamical quantities can be defined.

The problem of discontinuities is not settled with this statement. In a system with n molecules the partition function $P(n)$ is analytic, but $\lim_{n \rightarrow \infty} [P(n)]^{1/n}$ may be non-analytic. In this case the specific heat per molecule or some of its derivatives, though finite, may rise over any finite limit, if the number of molecules increases sufficiently. An example for this is provided by the gradient of specific heat in the theory of Bragg and Williams. As the number of molecules is large in macroscopic systems, experiment would give a discontinuity.

Discontinuities of this kind are not excluded by theorem I; but in special systems they may be excluded by a corollary to this theorem.

(Ia) If a thermodynamical system consisting of n particles and having the partition function $P(n)$ can be divided into smaller systems of j particles ($n = gj$) so that at a constant and finite value of j , $\lim_{g \rightarrow \infty} P(n) = [P(j)]^g$, the mean energy per particle and all its derivatives must tend to a finite limit in the $\lim_{n \rightarrow \infty}$.

This follows from the fact that in such a system the mean energy per particle and all its derivatives are equal to the mean energy per particle and its respective derivatives in a system of j particles, in which system all discontinuities are excluded by theorem I.

5. PROOF OF CONTINUITY OF C_v

An argument for the continuity of the C_v curve of β -brass is found as follows: The lattice is divided into cubes and the links are classified in links within the cubes and surface links between lattice points belonging to

different cubes. If the cubes are large enough, the number of surface links and their contribution to the energy may be neglected. In this approximation the cubes are statistically independent and the mean energy per atom, being independent of the number of cubes, is according to theorem I an analytic function of temperature.

In order to give a rigorous proof of continuity the cubes must be shown to become statistically independent in the limit of an infinite lattice, even if the energy of surface links is not to be neglected; moreover, the exchange of atoms of different cubes must be taken into account.

At first all possible numbers of A - and B -atoms are allowed in the lattice, so that every lattice point can be occupied by an A or a B atom, independently of the atoms by which the other lattice points are occupied.

It is convenient to introduce η_l , the energy of the link with the index l , as variable. η_l can take two values 0 and ϕ . In the one-dimensional chain any of these variables is allowed to assume its two values independently of the others. In the two- or three-dimensional lattice they are no longer independent. Of the links, which form a closed contour, an even number only is allowed to take the value 0 or ϕ . (With 4 links which form a square for instance, 0 or 2 or 4 links may have the energy 0 but not 1 or 3). If 2 (or 3) links are considered any of them can assume the energy 0 or ϕ whatever may be the energy of the other link (or other two links). The fluctuations of energy of different links are therefore independent:

$$\overline{[\sum_l \eta_l^2 - (\sum_l \eta_l)^2]} = \sum_l (\overline{\eta_l^2} - \bar{\eta}_l^2). \quad (11)$$

The lattice of $n = gj$ lattice points is divided into g cubes each containing j lattice points. The energy E_m of any configuration may be written as the sum

$$E_m = e_m + \epsilon_m$$

of e_m , the energy of the virtually separated cubes and ϵ_m the energy of the surfaces. The partition function

$$P = \sum_m e^{-E_m \tau} w_m = \sum_m e^{-(e_m + \epsilon_m) \tau} w_m$$

is enclosed between two limits as follows:

$$e^{-\mu \tau} < \frac{P}{\sum_m e^{-e_m \tau} w_m} < e^{-\lambda \tau},$$

if $\mu \geq \epsilon_m \geq \lambda$. The limits

$$\mu = \bar{e} + (\tfrac{1}{2}) (\bar{e}^2 - \bar{e}^2)^{\frac{1}{2}}, \quad \lambda = \bar{e} - (\tfrac{1}{2}) (\bar{e}^2 - \bar{e}^2)^{\frac{1}{2}}$$

are sufficiently wide.

The following estimate of the mean square of deviation shows that the overwhelming majority of configurations has surface energies lying within these limits.

By means of equ. (11) the relation

$$\bar{\epsilon}^2 - \bar{\epsilon}^2 = 3gj^{\frac{2}{3}}(\bar{\eta}^2 - \bar{\eta}^2) = \frac{3gj^{\frac{2}{3}}(\bar{E}^2 - \bar{E}^2)}{n} = \frac{3gj^{\frac{2}{3}}kT^2C_v}{L}$$

is obtained and, taking an experimental upper limit for the specific heat, $(C_v/L) < 8k$ the limits of the surface energy are estimated:

$$0 \leq \mu - \epsilon = \bar{\epsilon} - \lambda \leq 6^{\frac{1}{3}}kTg^{\frac{1}{3}}j^{\frac{1}{3}}.$$

The mean square of deviation appears to have the same magnitude as in independent systems.

The free energy of the surface links per atom is enclosed between two limits

$$(\bar{\epsilon}/n) + 6^{\frac{1}{3}} \frac{kT}{g^{\frac{1}{3}}j^{\frac{1}{3}}} > \frac{kT}{n} [\log \sum_m e^{-e_m \tau} w_m - \log \sum_m e^{-(e_m + \epsilon_m) \tau} w_m] > (\bar{\epsilon}/n) - 6^{\frac{1}{3}} \frac{kT}{g^{\frac{1}{3}}j^{\frac{1}{3}}}.$$

The free energy per atom of the infinite crystal ($\lim g \rightarrow \infty$) is

$$(1/n) F(n) = (1/j) F(j) + \bar{\epsilon}/n,$$

the sum of contributions of independent cubes and of the total surface; the latter contribution depends on temperature but not on the configurations of the cubes, and is accordingly the sum of contributions of independent surfaces attached to each cube:

$$\bar{\epsilon}(n) = g\bar{\epsilon}(j).$$

The partition function can be written

$$P(n) = e^{-\tau(n/j)[F(j) + \bar{\epsilon}(j)]} = [P(j)]^g \cdot [e^{-\tau\bar{\epsilon}(j)}]^g$$

as the product of two functions to each of which theorem Ia applies. The mean energy per atom and its derivatives are therefore continuous functions of the temperature.

So far the number of A and B atoms have been treated as variables. Actually these numbers are variable only in the single cubes but are constant in the lattice. This inconsistency has no effect upon the final result. A general theorem of statistical mechanics says that any quantity, being independent in the systems of an assembly and subject to a law of conservation, must appear as exponent of a canonical distribution with a modul which depends on temperature only (Delbrück and Molière 1936).

If the number of A and B atoms in the lattice is variable the partition function may be written

$$P(n) = \left[\sum_{\nu_\rho=0}^j p(j, \nu_\rho) \right]^g, \quad (12)$$

where ρ is an index attached to every cube ($1 \leq \rho \leq g$) and ν_ρ is the number of A atoms in the cube ($0 \leq \nu_\rho \leq j$); $p(j, \nu_\rho)$ is the partial sum in the partition function taken over all configurations with the same ν_ρ .

If the number of A atoms is $\frac{1}{2}n$, the partition function must be modified:

$$P(n) = \left[\sum_{\nu_\rho=0}^j p(j, \nu_\rho) e^{-\nu_\rho f(\tau)} \right]^g, \quad (13)$$

where $f(\tau)$ is determined by the condition:

$$j/2 = \frac{\sum_{\nu_\rho=0}^j \nu_\rho p(j, \nu_\rho) e^{-\nu_\rho f(\tau)}}{\sum_{\nu_\rho=0}^j p(j, \nu_\rho) e^{-\nu_\rho f(\tau)}}. \quad (14)$$

As the problem is symmetric in the numbers of A and of B atoms,

$$p(j, \nu_\rho) = p[j, (j - \nu_\rho)].$$

Equ. (14) can therefore only be satisfied if $f(\tau) = 0$ for all temperatures. The partition functions (12) and (13) are seen to be equal.

This proof shows, first, that Kirkwood's approximation is insufficient at the critical temperature, secondly, that Bethe's treatment of the one-dimensional chain is rigorously correct, but that his treatment of the two- and three-dimensional lattices is wrong at the critical temperature.

In § 3 the slope of the C_v curve was found to be finite; this conclusion is now confirmed.

6. THE SLOPE OF THE C_v CURVE

The partition function of neighbouring cubes can generally not be written as a product. It contains "coupling" terms, which remove the statistical independence of the cubes. In an infinite lattice the contribution of the "coupling" terms can be summarized as contributions of independent surfaces attached to every cube; this holds for any size of the cubes as shown in § 5. It follows that even a small number of cubes approaches statistical independence if only the number of atoms within the cubes is large enough.

The relative contribution of the "coupling" terms to this partition function gives a rough estimate for the approach to the statistical independence of the cubes. The contribution is diminished if their size increases.

In the semi-empirical method of § 3 the lattice is divided into aggregates of j atoms. The assumption is implied that these aggregates are statistically independent, and subdivisions of them are "coupled" by an apparent long distance interaction. This approximation involves a relation between the contributions of the "coupling" terms and the size of the cubes: at the critical temperature this contribution must be considerable for cubes of less than j atoms and decrease rapidly if the size of the cubes rises over j .

It will be shown in the following section that the empirical method can be justified on theoretical grounds. The partition function of neighbouring cubes is calculated. It will be shown that such a number of atoms in the cube j_0 exists, that for cubes smaller than j_0 the contribution of "coupling" terms is appreciable and becomes very small for cubes with more atoms. The value of j_0 will be seen to be of the same magnitude as found in § 3.

The partition function of neighbouring cubes may be written:

$$P = \sum_{E_1} \sum_{E_2} g_1(E_1) g_2(E_2) g_{12}(E_1, E_2),$$

where g_1 and g_2 are the partition functions of the virtually separated cubes 1 and 2 and g_{12} is a sum taken over all configurations of the surface links compatible with energies E_1 and E_2 of the cubes. This partition function is approximated by taking the sum g_{12} over such configurations only, which are compatible with an energy of either cube equal to their mean energy \bar{E} :

$$P = P_1 P_2 P_{12}(\bar{E}, \tau).$$

The functions P_1 and P_2 are assumed to be calculated according to the method of § 3. In deriving the partition function P_{12} the lattice points on the surface are classified in a and b points. It is convenient to define the a points in such a way as to put an a point on the surface of one cube opposite to an a point of the other cube. The link between two opposite surface points has the energy 0 or ϕ for atoms of the same kind or unequal atoms respectively. m denotes the number of lattice points on the surface of each cube. The calculation is simplified by the assumption that the number of A atoms on the surface is $\frac{1}{2}m$ and that the number of A atoms on a points of the surface is $q = \bar{p}m/j$, where p has the same significance as in § 3. By introducing the variable x of § 3 according to $\bar{p} = (1 + 2x)$,

$$q = \frac{1}{4}(1 + 2x)m$$

is obtained. As x is a function of \bar{E} ,

$$\bar{E} = j\phi(\frac{3}{2} - 6x^2),$$

q is also a function of \bar{E} ,

$$q = q(\bar{E}) = [1 + 2(\frac{1}{4} - \bar{E}/6\phi j)^{\frac{1}{2}}].$$

In deriving the partition function the configurations on the surfaces are to be enumerated, which are compatible with given energies of the cubes. If r and s are the numbers of a and b points respectively, the links of which are occupied by pairs of A atoms, the number of configurations with given q , r , s is found to be

$$\binom{q}{r} \binom{\frac{1}{2}m - q}{s},$$

the corresponding energy of the links being $2\phi(r+s)$ and

$$0 \leq r \leq q \quad (\frac{1}{2}m - 2q) \leq s \leq (\frac{1}{2}m - q),$$

provided that $q < \frac{1}{4}m$. (The terms with $q > \frac{1}{4}m$ would have a sum equal to the sum of the terms taken into account and may be neglected in the following discussion.)

The partition function

$$\begin{aligned} P_{12} &= \sum_{r=0}^q \sum_{s=\frac{1}{2}m-2q}^{\frac{1}{2}m-q} \binom{q}{r} \binom{\frac{1}{2}m - q}{s} e^{-2(r+s)\phi\tau} \\ &= (1 + e^{-2\phi\tau})^{\frac{1}{2}m} - (1 + e^{-2\phi\tau})^q \sum_{s=0}^{\frac{1}{2}m-2q} \binom{\frac{1}{2}m - q}{s} e^{-2s\phi\tau} \end{aligned}$$

can be written

$$P_{12} = Q'(\tau) - Q''(\tau, \bar{E}),$$

Q'' being the contribution of the "coupling" terms. As the partition function is equal to

$$P_{12} = (1 + e^{-2\phi\tau})^q \left[\sum_{s=0}^{\frac{1}{2}m-q} \binom{\frac{1}{2}m - q}{s} e^{-2\phi\tau s} - \sum_{s=0}^{\frac{1}{2}m-2q} \binom{\frac{1}{2}m - q}{s} e^{-2\phi\tau s} \right],$$

the relative magnitude of Q'' and Q' is readily estimated. If the maximum term of the first sum is included in the second sum, both sums are of equal magnitude; if the maximum term is shifted beyond the limits of the second sum, the latter becomes small compared with the first.

The maximum term is given by

$$\binom{\frac{1}{2}m - q}{\bar{s}} e^{-2\phi\tau \bar{s}},$$

where

$$\bar{s} = (\frac{1}{2}m - q) e^{-2\phi\tau} / (1 + e^{-2\phi\tau}).$$

It appears in the second sum if $\bar{s} \leq (\frac{1}{2}m - 2q)$. The relative magnitude of Q' and Q'' changes rapidly for $\bar{s} = \frac{1}{2}m - 2q$. All these numbers are functions of j, m according to

$$m = 3j^{\frac{1}{3}}.$$

The factor 3 is introduced to account for the interaction of a cube with all its neighbours. x depends on j according to

$$x^2 = \frac{1}{4} - \bar{E}/6j,$$

and (at the critical temperature) according to equ. (4a)

$$\bar{E} = j\phi(\frac{3}{2} - 3(3/j\pi)^{\frac{1}{3}}).$$

It follows that

$$x = -(3/4\pi j)^{\frac{1}{2}} = -0.70j^{-\frac{1}{2}}, \quad (15)$$

$$q = (\frac{3}{4})j^{\frac{1}{3}}(1 - 1.40j^{-\frac{1}{3}}), \quad (16)$$

$$(\frac{1}{2}m - q) = (\frac{3}{4})j^{\frac{1}{3}}(1 + 1.40j^{-\frac{1}{3}}), \quad (17)$$

$$(\frac{1}{2}m - 2q) = 2.10j^{\frac{1}{3}}, \quad (18)$$

and as at the critical temperature $e^{-2\phi\tau}/(1 + e^{-2\phi\tau}) = 0.339$,

$$\bar{s} = 0.254j^{\frac{1}{3}}(1 + 1.40j^{-\frac{1}{3}}). \quad (19)$$

The value of j for which the maximum term of the first sum is equal to the upper limit of the second sum

$$\bar{s} = \frac{1}{2}m - 2q \quad (20)$$

is calculated according to equ. (17), (18) and (19) from the equation

$$0.254j^{\frac{1}{3}}(1 + 1.40j^{-\frac{1}{3}}) = 2.10j^{\frac{1}{3}}, \quad (20a)$$

the solution of which is $j_0 = 2.2 \times 10^3$.

For $j < j_0$ the ratio $Q''/Q' \approx 1$, the difference $Q' - Q''$ being of the magnitude $m^{-\frac{1}{2}}$. For $j > j_0$ the ratio decreases exponentially. An upper limit of the ratio for $j = 10^4$ is $0.6 \times e^{-5}$, the actual value is likely to be much smaller. Such a relation between Q''/Q' and j at the critical temperature is consistent with the concept implied in the method of § 3. The numerical value of j_0 is of the same magnitude as 10^4 the number of atoms within one aggregate, which number was found in § 3 in order to yield the experimental gradient of specific heat at the critical temperature.

The value of j_0 increases if the temperature is diminished. Near $T = 0$, $q = 0$, $s = 0$, $\frac{1}{2}m - 2q = \frac{1}{2}m$ for any value of j so that equ. (20a) has no finite solution. This is not in contradiction with the result of § 5, for in the limit

of an infinite lattice the partition function has the form of a power of the partition function of one pair of atoms.

The interpretation of experimental facts as given in § 3 appears to be compatible with the molecular model.

In conclusion, the author wishes to express his thanks to Sir William Bragg and the managers of the Royal Institution for the excellent opportunities given to him for carrying out this work at the Davy Faraday Laboratory. He is also much indebted to Dr A. Müller for many interesting and helpful discussions.

SUMMARY

A theoretical investigation is made of the specific heat of β -brass. The following assumptions are made in the course of the calculation:—

The interaction energies of the atoms vary with the square of change of volume.

The energy minimum of the average of Cu-Cu and Zn-Zn energies are lying at higher volume than the minimum of Cu-Zn.

The partition function at constant volume is calculated according to the theory of Bragg and Williams; the corresponding partition function at constant pressure is obtained by taking an average over the variable volume.

It is shown that the specific heat curve can be adjusted so as to give good agreement with experiment for temperatures below the maximum of specific heat.

The theory of Bragg and Williams is applied to a lattice containing a finite number of atoms. By introducing this number agreement is obtained with the experimental slope of the C_p curve beyond the maximum. A theoretical estimate is made which justifies the introduction of this finite number.

A general theorem of statistical mechanics is derived, according to which all thermodynamical quantities in physically realizable systems are analytic functions of the temperature.

REFERENCES

- Bragg and Williams 1934 *Proc. Roy. Soc. A*, **145**, 699.
Bethe 1935 *Proc. Roy. Soc. A*, **150**, 552.
Delbrück and Molière 1936 *Abh. preuss. Akad. Wiss.* no. 1.
Kirkwood 1938 *J. Chem. Phys.* **6**, 70.
Lennard-Jones and Devonshire 1937 *Proc. Roy. Soc. A*, **163**, 53.

- Moser 1936 *Phys. Z.* **37**, 737.
Mott and Potter 1937 *Nature, Lond.*, **139**, 411.
Müller 1938 *Proc. Roy. Soc. A*, **166**, 316.
Nix and Shockley 1938 *Rev. Mod. Phys.* **10**, 1.
Opechowsky 1937 *Physica*, **4**, 181.
Steinwehr and Schulze 1934 *Phys. Z.* **35**, 385.
Sykes 1935 *Proc. Roy. Soc. A*, **148**, 422.
-

A further study of the problem of nuclear isomerism:
the application of the method of coincidence counting
to the investigation of the γ -rays emitted by uranium
Z and the radioactive silver Ag^{106}

BY N. FEATHER, PH.D., *Fellow of Trinity College, Cambridge*, AND
J. V. DUNWORTH, B.A., *Denman Baynes Research Student, Clare College*

(Communicated by E. V. Appleton, F.R.S.—Received 3 August 1938)

1. INTRODUCTION

The application of the method of coincidence counting by electrical means to the study of essentially nuclear phenomena was made first by Bothe and v. Baeyer in 1935.* Since that time, and chiefly as the result of further experiments carried out in Heidelberg, much important information has been obtained by the use of this method. It is the object of the present paper to describe the application of the method to the elucidation of certain features of the problem of isomerism in radioactive nuclei. As pointed out in a discussion of this problem by Feather and Bretscher (1938),† if such isomerism is to be explained in terms of close nuclear states differing considerably in nuclear spin, it is frequently to be expected that the disintegration of one or other isomeric nucleus will be followed by the emission of two or more quanta of γ -radiation in succession. The aim of the coincidence

* The use of coincidence counting by visual means is much older than this. In 1910 Geiger and Marsden postulated the successive emission within a small fraction of a second of two α -particles from the same atom, after counting "coincident" scintillations, and 22 years later Cockcroft and Walton (1932) employed precisely the same method in preliminary tests concerning the simultaneous projection of two α -particles in the single act of disintegration of a lithium nucleus by a proton.

† Referred to as Paper I in what follows.

experiments, therefore, was to explore this possibility in a number of cases—and particularly to look for evidence of paired γ -rays following the β -disintegration of uranium Z, as suggested in Paper I. Clear evidence of the latter effect has been found, and the case of isomerism in radioactive silver (Ag^{106}) has also been studied with positive results. Since the information provided by the coincidence method cannot profitably be discussed without data concerning the particle radiations emitted by the isomeric nuclei as well, full details of absorption measurements on the β - and γ -radiations from the radioactive silver are also given in this paper. During the investigation one of us (N. F.) has been responsible chiefly for the absorption measurements, using a single counter, the other (J. V. D.) for the coincidence experiments with two counters. The artificially radioactive sources were prepared for us in the High Voltage Laboratory and we are greatly indebted to the workers in that Laboratory, and particularly to Dr W. E. Burcham, for their efficient and kind co-operation in this undertaking.

2. THE METHOD OF COINCIDENCE COUNTING

Let us suppose that a radioactive source which emits γ -radiation to the extent of n_1 quanta of energy E_1 , n_2 quanta of energy E_2 , ..., n_r quanta of energy E_r , per second, is brought to a position near a thick-walled tube counter subtending a solid angle $4\pi\alpha$ at the source. Let the efficiency of the counter for each of the several quantum radiations be $\epsilon_1, \epsilon_2, \dots, \epsilon_r$, and ν impulses per second be the natural background count of the instrument. Then, if none of the particle radiations from the source enters the counter,

$$N = \alpha \sum_{k=1}^{k=r} n_k \epsilon_k + \nu$$

impulses per second is the average counting rate in the presence of the source. If a second exactly similar counter be now placed so as also to subtend a solid angle $4\pi\alpha$ at the source, a coincidence rate of

$$A = 2N^2\tau + c \quad (1)$$

coincidences per second will be observed, on account of the finite resolving time (τ sec.) of the coincidence circuit and because of penetrating radiation effects (c coincidences per second) alone. Further coincidences may be recorded if certain of the γ -radiations are time-correlated (in that two or more quanta are emitted in rapid succession from the same nucleus: the effect which we are chiefly investigating), or because of the production of

annihilation radiation in the material immediately surrounding the source*—either because the main particle radiation consists of positive electrons (when this coincidence effect will be considerable), or, as a relatively infrequent result, during the absorption of primary β -particles or of γ -rays emitted by the source. Since, in any case, the annihilation radiation coincidence rate must be proportional to the strength of the source, we may write $\beta(N - \nu)$ impulses per second to represent this effect (β being known only when full information concerning the radiations from the source is available, and the geometrical disposition of the absorbing material is taken into account)—and thus obtain as the total number of coincidences per second

$$B = \beta(N - \nu) + 2N^2\tau + c \quad (2)$$

—in the absence of time-correlation in the γ -radiation from the source.

To calculate this last remaining effect, particular assumptions must be made. Let us first suppose that as the result of particle disintegration the nucleus is always left in a given state of excitation and that the process of de-excitation follows an invariable sequence in which r separate quanta are emitted. Then, in the above notation, $n_1 = n_2 = \dots = n_r = n$, and a coincidence rate of

$$C = n\alpha^2 \sum_k \sum_l \epsilon_k \epsilon_l (\neq k) \quad (3)$$

per second is to be expected.† It is instructive to compare this “true” coincidence rate with $(N - \nu)$ the number of impulses per second in a single counter due to the γ -radiation. If R represents the ratio of these two rates of counting, we have

$$R = \alpha \Sigma \Sigma \epsilon_k \epsilon_l / \Sigma \epsilon_k \quad (4)$$

in the particular case considered. For sake of brevity this result may be written

$$R = (r - 1)\alpha\bar{\epsilon}, \quad (5)$$

where $\bar{\epsilon}$ is a representative efficiency of the counter for γ -radiation of the general hardness in question. Since, with complex radiation, a rough value of $\bar{\epsilon}$ may often be evaluated from considerations of the general absorbability of the radiation as a whole, (5) is sometimes useful as indicating a value for r , the number of quanta emitted in cascade‡ following a given instance of particle disintegration.

* It is here assumed, as in fact was the case in the experiments to be described, that the source is placed symmetrically between the counters.

† First order terms, only, are included.

‡ This result is essentially that used by Fleischmann (1936) and also by Aoki (1937) for investigations of the capture radiation emitted in processes of slow neutron capture by nuclei.

Clearly, R may have a value different from that given by (4) or (5) whenever the previous particle disintegration results in nuclei in different states of excitation on different occasions, or when different modes of de-excitation of a single nuclear state are in competition one with another. In the simplest case, in which de-excitation gives rise to quanta of energies E_1 , E_2 and E_3 with relative probabilities $n_1 : n : n$ (such that quanta of the last two energies alone are time-correlated, and these completely so), simple calculation leads to the following result (Maier-Leibnitz 1936):

$$R = \frac{2n\alpha\epsilon_2\epsilon_3}{n_1\epsilon_1 + n(\epsilon_2 + \epsilon_3)}. \quad (6)$$

It is the general aim of the coincidence experiments, therefore, to separate out the contributions to the total coincidence rate according to the scheme suggested by (1), (2) and (3) above, and to interpret the "true" coincidence rate, indicating time-correlation in quantum emission, in terms of (4), (5) and (6), and similar appropriate results. As already indicated, the final validity of any such interpretation obviously depends to a large extent on the nature of our knowledge concerning the energies and relative intensities of the particle and quantum radiations from the elements examined: without this knowledge it is clearly impossible to substitute the requisite numerical values in the expressions concerned.* With this remark, therefore, we pass to a description of the apparatus and the presentation of the experimental results.

3. THE APPARATUS AND ITS CALIBRATION

Experimental arrangement. The geometrical arrangement of tube counters which was adopted as standard throughout the experiments is shown diagrammatically in fig. 1. The cylindrical brass walls of the counters were 5 cm. long, 2 cm. in diameter and 0.58 mm. thick. The counters were placed

* The final accuracy of the numerical result—as distinct from its interpretation—may be investigated simply as follows. Let the total time of observation with a source in position be T sec. and suppose that a similar time has been spent in a subsidiary experiment with another source in determining the number of chance and "cosmic" coincidences. Then the probable error in the number of true coincidences per second is of the order of $(B + C + A)^{\frac{1}{2}}T^{-\frac{1}{2}}$. With T constant, the latter quantity decreases as the strength of source increases, until, when $2N^2\tau$ becomes the dominant term in B (and A), it tends to the constant value $2(\tau/T)^{\frac{1}{2}}R^{-\frac{1}{2}}$. For sufficiently strong sources, therefore, the full benefit of a small resolving time may be utilized. In our case, with $R \sim 10^{-3}$ and $\tau \sim 10^{-5}$ sec., 1% accuracy will be seen to necessitate roughly 100 hr. counting, both on the chance and true coincidence rates, using sources of adequate strength. One hour's counting on each might be expected to give an accuracy of 10%.

with their axes parallel, at 3 cm. separation, in the same horizontal plane. In this way only 1 cm. was left between the nearest points of the counters and it was found necessary to insert an earthed screen between them in order to prevent the discharge of one counter influencing the electrical circuit connected to the other. For sake of symmetry two such thin metal screens were inserted as shown. Moreover, these screens served to support a V-shaped source holder of the same thin metal (~ 0.44 g./cm.²). It was found in practice that the use of this holder enabled the source to be removed and replaced in the symmetrical position with an accuracy represented by $\pm 2\%$ in the counting rate of either counter. The arrangement is very similar to that of Maier-Leibnitz (1936) except that no tungsten (or lead) was used. Maier-Leibnitz worked with tungsten between the counters to reduce the effect of annihilation radiation (represented by the first term in (2)): in respect of the conditions of our main experiments calculation based upon available data (cf. Feather and Dunworth 1938*a*) showed that this effect was negligibly small.* The "natural" coincidence rate due to penetrating radiation effects was found to be about 12.5 per hr. with the standard arrangement; with the counter axes in the same vertical (rather than horizontal) plane, however, this rate was about 8 times greater.

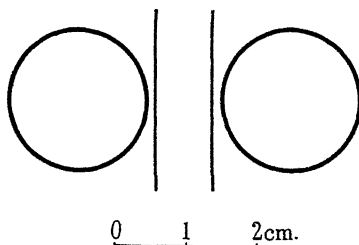


FIG. 1

The electrical connexions were as follows. Each counter was connected in a modified† Neher-Harper circuit (Neher and Harper 1936) which was followed by pulse-sharpening stages. Finally, the two circuits were combined by the usual Rossi method (Rossi 1930). This arrangement was normally used with a resolving time of 6×10^{-6} sec. (10^{-7} min.). It is important to note that the satisfactory use of this small resolving time requires that every electrical pulse reaches the end of the appropriate valve line within a

* It can be shown that the occurrence of annihilation radiation in the general radiation from a source will result in an experimental value of R (p. 568), which is greater or less than the true value referring to the γ -radiation alone, depending upon whether the actual value of R is less or greater than about 1.2×10^{-3} (cf. p. 573).

† Modifications were incorporated due to Dr W. B. Lewis, of this Laboratory.

time rather less than 10^{-7} min. from the time of emission of the γ -ray quantum responsible for actuating the counter. Thus not only must the delay in the intermediate electrical stages be reduced below this value, but also the response of the counter must be "instantaneous" within these limits (cf. Werner 1934). In respect of the latter condition all that can be said is that it appears to have been satisfied with the counters which we used: no increase in the coincidence counting rate—over and above that due to chance coincidences—(i.e. no increase in the "true" coincidence rate) was recorded when the resolving time was raised to 5×10^{-4} sec. The counters were filled with argon (~ 6.5 cm. pressure) and ethyl alcohol vapour (1.5 cm. pressure).

Observations on the voltage-dependence of the single counting rates of the two counters showed that the characteristic curves were horizontal* from a few volts to well above 100 V above the starting potential. The counters were generally operated at 80–100 V overvoltage. Further, there was no appreciable loss of single counts in either counter at the highest rates used. This is important, since any such loss would result in a low value for the experimentally determined ratio R . In some experiments sources were used with which single counting rates of 6000 min.^{-1} were reached. As already pointed out (p. 569), it is necessary to employ such sources in order to derive full benefit from the small resolving time of the coincidence circuit.

Calibration. In the present connexion calibration of the counters is intended to refer to the determination of the variation of counter efficiency with the quantum energy of the incident radiation. Clearly, full information on this subject is necessary before equations (4), (5) and (6)—and similar results—can be applied in the interpretation of observational data. In respect of the equations, however, one point should be noted before we proceed. Throughout the theoretical discussion of §2 the solid angle ($4\pi\alpha$) and the efficiency of the counter (ϵ) have been assumed to be separable factors determining jointly the probability of response of the counter. This, in fact, is an over-simplification of the position. With thick-walled counters the probability of response obviously depends upon the point of origin of the recoil- or photo-electron in the wall of the counter, and an over-all response factor of $\int \epsilon d\alpha$ should properly replace the product

* Klemperer (1934) has used coincidence counters in the form of half cylinders as shown: □□. Whilst this may ensure a slightly greater solid angle, any such advantage is certainly offset by the absence of an appreciable horizontal portion on the characteristic curve of the counters.

$\alpha\epsilon$ in all our equations. Calibration, therefore, refers to the determination of this response factor, $(\alpha\epsilon)$, in relation to quantum energy. The experiments which were made for the purpose of such calibration are listed below.

(1) Experiments requiring a knowledge of the energies and intensities of the important quantum radiations from thorium active deposit (Oppenheimer 1936; Ellis 1938): (a) Determination of the counting rate in a single counter when a source of known strength, consisting of Th (B + C) in equilibrium, was placed in the standard position between the counters. (b) Determination of the single counting rate as a function of the time for a short exposure (~ 1 sec.) deposit of thorium active deposit. (c) Determination of the relative coincidence rate, R (p. 568), with a source of Th (B + C) in equilibrium. (d) Investigation of the β - γ coincidences with a similar source, i.e. the coincidences between the impulses in a counter with a thin window, due to the primary β -particles from the source, and the impulses in a thick-walled counter due to the γ -rays. The radiations chiefly involved in these experiments have quantum energies 0.24×10^6 eV (Th B), 0.51, 0.58 and 2.62×10^6 eV (Th C").

To a first approximation (adjusted eventually in view of the results of all four experiments) experiment 1 (a) gives the response factor $(\alpha\epsilon)$ for the hardest of these radiations. Then the ratio of the initial to the final counting rate in 1 (b) gives $(\alpha\epsilon)$ for the softest radiation.* Experiment 1 (c), if the level scheme of Oppenheimer (1936) is correct, gives $(\overline{\alpha\epsilon})$ for the mixed radiation of 0.51 and 0.58×10^6 eV energy, on the basis of (4). Finally, with a similar proviso concerning the level scheme, 1 (d) gives the sum of the response factors for the hard radiation and the mixed radiation just mentioned. In respect of 1 (c) and 1 (d) it should be emphasized that these coincidence experiments provide the first objective test of the correctness of this level scheme—and so of the original suggestion of Ellis and Mott (1933) that excitation of the final nucleus in the transition Th C". Th Pb occurs in almost every disintegration in a level of 3.2×10^6 eV energy. It may be said at once that in each case coincidences were found—and roughly with the frequency which was expected: the value of $(\overline{\alpha\epsilon})$ for the softer radiations of Th C" deduced from 1 (c) and 1 (d) is clearly of the correct order of magnitude. This conclusion depends upon 240 min. counting of

* Actually, the counting rate in 1 (b) *decreased* initially before following the normal rise to be expected for a source collected by short exposure to thorium emanation. Eventually, however, the initial decrease was satisfactorily explained as due to Th C" collected by recoil from the walls of the exposure vessel, and could therefore be corrected for. This explanation naturally suggested itself when a complete analysis of the time variation of the counting rate indicated a half-value period of 3.2 min. for the initial effect.

coincidences with two sources in experiment 1 (c), during which time 1428 coincidences were recorded, and 571 min. counting in experiment 1 (d) yielding 2072 coincidences.

(2) An experiment with annihilation radiation (quantum energy 0.51×10^6 eV).

Roughly 200 coincidences due to the annihilation radiation from the positrons emitted by two short activation sources of Ag^{106} (half-value period 24.5 min.) were recorded and the relative coincidence rate, R , determined by frequent counts with a single counter during the decay of the activity. As pointed out by Klemperer (1934), in discussing a similar experiment in which the positron-active C^{11} was employed,* if the quanta in this case are emitted strictly in opposite directions, then, for a point source symmetrically situated between identical counters, $R = \epsilon$. For our present purpose, therefore, it is necessary to make a separate determination of α before we can obtain the appropriate point on the $(\alpha\epsilon)$ -energy curve, as required. Since, however, α and ϵ are not, as here supposed, separable factors (p. 571), the matter is again further complicated: the determination of R in the coincidence experiment with annihilation radiation gives

$\int \epsilon^2 d\alpha / \int \epsilon d\alpha$, whereas, if we are to combine a measurement of efficiency with a separate determination of the geometrical solid angle, $4\pi\alpha$, the efficiency which we require is $\int \epsilon d\alpha / \int d\alpha$ —always a smaller quantity. On the other

hand, since any actual arrangement must deviate appreciably from the ideal of a point source symmetrically placed between identical counters, another error is necessarily introduced in the opposite sense. Clearly, the annihilation radiation experiment can at best lend support to a value of $(\alpha\epsilon)$ determined from other considerations. Actually, in the present experiment with the positrons of Ag^{106} , we found $R = (1.18 \pm 0.08) \times 10^{-3}$, pointing to a value of $(\alpha\epsilon)$ of the order of 0.3×10^{-3} for radiation of quantum energy 0.51×10^6 eV. This is in reasonable agreement with our previous results.

In fig. 2 these results are given in graphical form and the attempt is made to construct a curve which for the present will best represent the values variously obtained. It is to be hoped that in the near future further experiments with other sources emitting radiations of different quantum energy will enable any necessary corrections to be made (for some information concerning the energy range 1 to 2×10^6 eV, see Feather and Dunworth

* Prior to the present experiment, coincidences due to positron annihilation had also been detected by Alichanian, Alichanow and Arzimovitch (1936)—using the radioactive phosphorus, P^{30} .

1938*b*). Many curves similar to that of fig. 2 have already been published. Some of these have been deduced mainly from experimental results, some chiefly on theoretical grounds (v. Droste 1936, 1937; Yukawa and Sakata 1937; Sizoo and Willemssen 1938). Before making detailed comparison with any of them, however, it must be remembered that they refer to thick-walled counters which in general subtend a small solid angle at the source; our curve, on the other hand, refers explicitly to the case in which the solid angle is large.

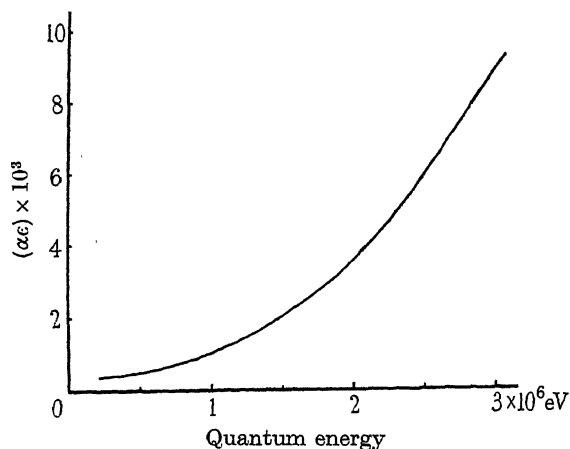


FIG. 2

Resolving time. In view of the doubts which have sometimes been expressed regarding the validity of equation (1), with τ constant, as a basis for calculating the correction for chance coincidences in an experiment like the present one, a considerable time was spent investigating this matter. Two chief considerations are involved. In the first place it is clearly a matter for investigation whether, when the impulses in the two counters are unquestionably without correlation, the behaviour of the coincidence circuit may simply be described by means of a constant resolving time, after the manner adopted in (1). In order to use this equation we must first show that this is actually the case.* The second question is whether, with a γ -ray source between two counters in close proximity, both may possibly be discharged as the result of interactions ascribable to a single γ -ray quantum. This question must receive a negative answer before equation (1) can be regarded as satisfactory.

* It would not be so, for example, if the form of the voltage pulse depended noticeably upon the counting rate, or was unduly sensitive to the voltage at which the counter was operated.

The first question has generally been investigated by counting coincidences with the two counters widely separated, each being irradiated by the γ -rays from a weak source. The observed coincidence rate is plotted against $N_1 N_2$, the product of the rates of counting in the two counters separately, and a straight line is obtained passing through the origin (c , the contribution to the coincidence rate due to penetrating radiation effects, is negligibly small in this arrangement). A similar experiment was carried out with our counters and the expected straight-line relation was obtained. It was also obtained (though the straight line no longer passed through the origin)

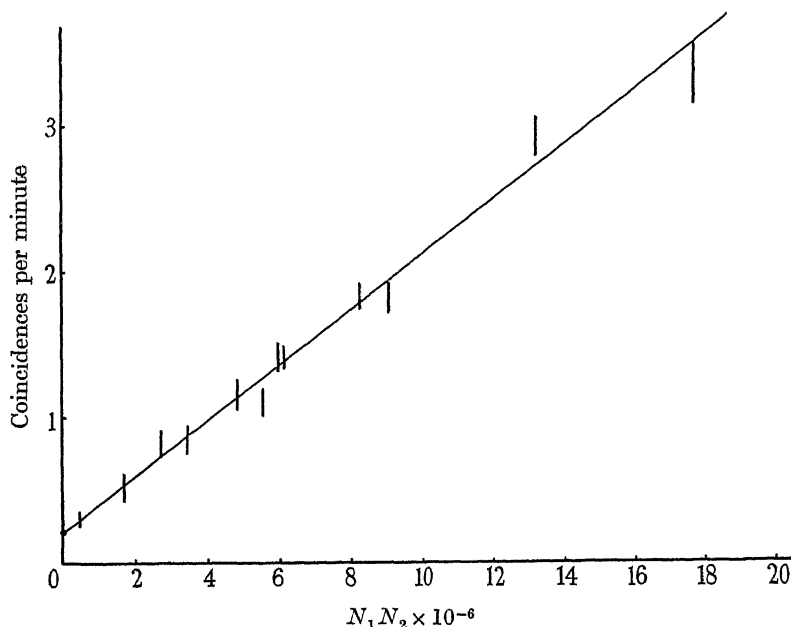


FIG. 3

when the experiment was repeated with the counters at the standard separation and a strong source of UX was so placed that the effect in each counter was chiefly due to the primary β -particles of UX_2 which had penetrated the 0.58 mm. wall of the counter. The results of these experiments are given in fig. 3. The slope of the straight line is the same as was obtained with the counters separated; in this case the intercept on the y -axis gives c , the coincidence rate due to penetrating radiation effects in the standard arrangement.

As regards the second question raised above, the coincidence rate was investigated when the counters were irradiated with the γ -rays from a

relatively strong source placed at a considerable distance from the counters—which again occupied the standard positions. With such an arrangement possible coincidences due to paired γ -rays from the source (a small sealed radium preparation was used) could be neglected on account of the small solid angle subtended by the counters at the source. Nevertheless, a coincidence rate over and above that given by the straight line of fig. 3 was obtained for some positions of the source. A survey of all the data appeared to show that this excess was noticeable only when there was a reasonable chance that a quantum scattered in producing a recoil electron in one counter should pass into and suffer scattering with recoil in the other. This effect, which might be expected to contribute a coincidence rate linear in N_1 and N_2 , was in fact found to be relatively most important at slow rates of counting. At any rate of counting, however, with the source situated in the symmetrical position midway between the counters, such a process as we have been considering would clearly be least effective; the maximum energy of the quantum scattered through 180° in the Compton effect is about 0.25×10^6 eV—and for such radiation the efficiency of registration is already very small. In fact it was finally concluded that no correction for this effect was necessary in the main experiments—and as a similar conclusion regarding the effect of pair production has already been reached (p. 570), it appears that the use of equation (1), with the constant value of τ deduced from fig. 3, is entirely justified.

4. EXPERIMENTAL RESULTS

Uranium Z. As already mentioned, the suggestion made in Paper I was that in the majority of cases the emission of the primary β -particle from the nucleus UZ is followed immediately by the emission of two quanta of γ -radiation. Absorption measurements indicated that these quanta were unlikely to differ much in energy and a mean energy of 0.7×10^6 eV was deduced for the radiation as a whole. From fig. 2 for such radiation we may write $(\overline{\alpha\epsilon}) = 0.65 \times 10^{-3}$ —and that this is also the appropriate value for R follows from the fact that we assume $r = 2$ in (5). The value found experimentally for R was $(0.85 \pm 0.06) \times 10^{-3}$ —indicating reasonable agreement with expectation. If any significance be attached to the fact that the experimental is greater than the “expected” value of R , it may be concluded that the γ -radiation from uranium Z very probably has a mean quantum energy somewhat in excess of 0.7×10^6 eV. It will be observed that this conclusion fits better with the level scheme of fig. 7, Paper I, than do the results of the

γ -ray measurements alone.* The experimental value for R was based chiefly on observations with three strong sources of uranium Z, represented in fig. 4. Here the residue of the observed coincidence rate, remaining after correction for chance coincidences and penetrating radiation effects, is plotted logarithmically against the time—and straight lines† are drawn

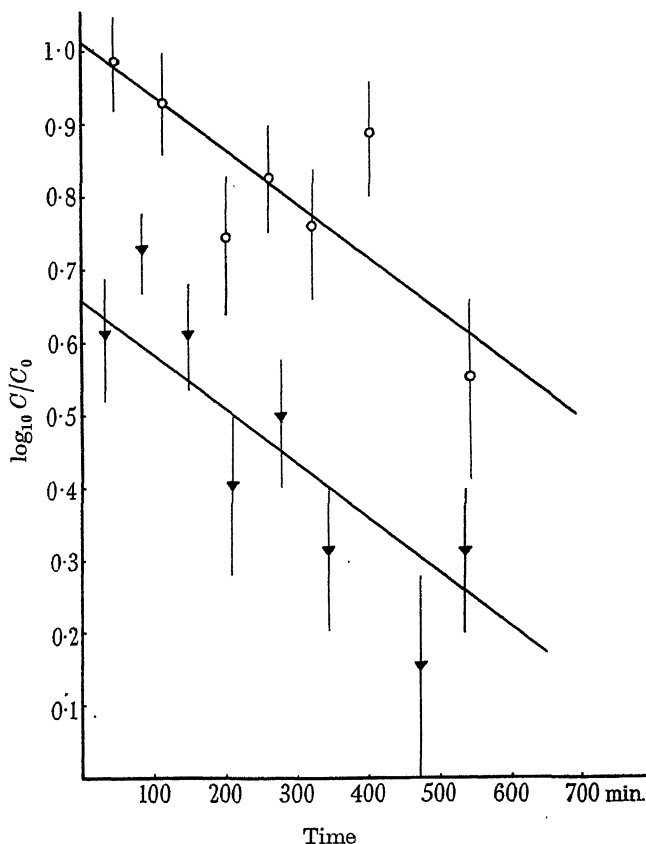


FIG. 4

amongst the points with the slope appropriate to the known rate of decay of uranium Z ($\tau = 6.7$ hr.). It is clear that the residual coincidence rate so calculated is proportional to the strength of the source. If this treatment of

* Although it is not completely excluded that the successive emission of three quanta of softer radiation, rather than two quanta of harder radiation, is the most usual happening.

† Only two such lines are drawn, since the observations with the two weaker sources have been combined.

the data had led to any other result, naturally our whole interpretation of the observed coincidences would fail.

Ag^{106} . The occurrence of nuclear isomerism in the case of Ag^{106} was suggested first by the observation of Pool, Cork and Thornton (1937) that a negative electron activity of half-value period about 8 days was obtained together with the already known positron activity of 24.5 min. period in each of four distinct methods of activation. The positron period had definitely been assigned to Ag^{106} by earlier workers (Bothe and Gentner 1937; Heyn 1937). In subsequent papers Pool (1938) and Pool and Campbell (1938) put forward additional support for this suggestion of isomerism. Moreover, because they found the long-period negative electron activity to be accompanied by an abnormally large amount of γ -radiation, they further suggested that transformation by K -electron capture (followed by γ -ray emission) might also occur with the long-lived species. Then the transformation $\text{Ag}^{106} \rightarrow \text{Pd}^{106}$ would be taking place with each isomer—and the possibility of checking the balance of energy (on the assumption that the final state of the product nucleus is the same in each case) would naturally present itself (see fig. 8). The experiments now to be described were designed to explore these and other assumptions in greater detail than had been done by the earlier workers.

The radiations from the short-lived isomer were investigated by bombarding a silver foil of 0.085 g./cm.² thickness for about an hour with the neutrons from lithium bombarded with deuterons (30 μA at 950 kV), starting measurements always half an hour after the end of the bombardment. About 3 cm.² of foil were used (usually at about 2 cm. distance from the mica window of the tube counter)—and the absorption experiment was carried out exactly as described in Paper I. Fig. 5*a* shows the results for absorber thicknesses including the particle end-point and fig. 5*b* the logarithmic absorption curve for the positrons alone. In constructing these curves correction has in all cases been made for the natural effect of the counter and also for the long-period activity which was not entirely negligible in amount, in spite of the shortness of the bombardment. For fig. 5*b* the annihilation radiation effect has been subtracted—after extrapolation of the “ γ -ray background” of the other curve (5*a*). From the general shape of the logarithmic curve it is at once clear that the spectrum of positrons is simple to a high degree: by a detailed comparison of this curve with the curves for the β -particles of radium E and uranium X₂ given in Paper I a positron end-point at 0.95 g./cm.² is deduced. On the basis of a slightly modified range-energy relation, we may therefore say that the positron spectrum (which is not complex) has an upper energy limit

at $(2.04 \pm 0.05) \times 10^6$ eV.* The value quoted by Pool and Campbell (1938) is 1.9×10^6 eV. The lack of complexity in the positron spectrum, taken together with the numerical result of the coincidence experiment already described (p. 573), suggests that the γ -ray background in fig. 5*a* is due

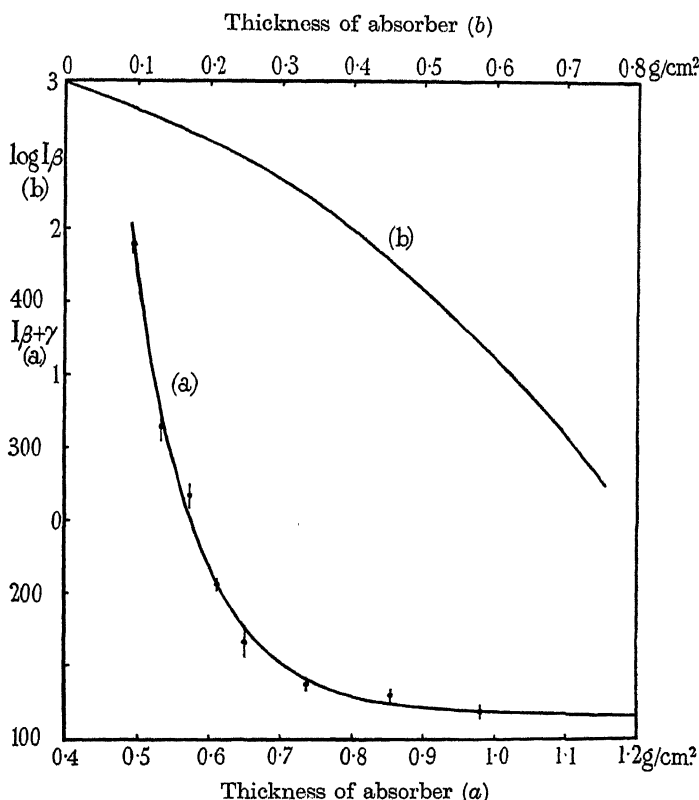


FIG. 5

almost entirely to annihilation radiation. A γ/β ratio of $1:77 \pm 2$ was deduced for this radiation, after the usual correction for absorption of the positrons in the material of the source (cf. Paper I, p. 540). Clearly, this ratio should be standard, for our experimental arrangement, for all positron-

* A further discussion of the possibilities of the absorption method (distinction between simple and complex spectra, together with other considerations) will be given by one of us (N. F.) at an early date. One subsidiary result of this discussion is that the original empirical range-energy relation, $R = 0.511E - 0.091$ (Feather 1930), should be replaced by $R = 0.543E - 0.160$ —in agreement with a recent suggestion of Widdowson and Champion (1938). No great changes ($> 0.09 \times 10^6$ eV for $0.7 < E < 3.0$) in accepted energy values follow from this alteration.

active substances which do not emit any appreciable amount of nuclear γ -radiation.

When the attempt was made to investigate the particle radiation from the long-lived isomer it became obvious at an early stage that a radioelement of half-value period intermediate between 24.5 min. and 8 days was produced by neutron bombardment. In general the bombardment lasted for about 8 hr., and thin sheets of silver (0.085 g./cm.^2) were activated, as before. The decay of activity of such a target showed periods of 12.3 hr. and about 8 days, roughly equally developed. Since no detectable amount of copper was present in the silver (the period of Cu^{64} is 12.8 hr.), it was concluded that radioactive palladium had been produced in a (n^1, H^1) transformation. It was known that a negative electron activity of 13 hr. period could be produced by bombarding palladium with deuterons (Krauss and Cork 1937)—and that this activity had tentatively been ascribed to Pd^{109} . Our identification appeared very reasonable in view of these facts, but apart from this the 12.3 hr. period was not further investigated. The radiations of the 8-day period were studied by the standard absorption method after allowing the supposed palladium activity to decay. Fig. 6 gives the results obtained with a thin silver foil in the usual position, showing at once the abnormally high γ/β ratio, as first reported by Pool (1938). On account of the rather low accuracy of individual points in fig. 6—and because of the very large “ γ -ray background”—it is not easy to make definite pronouncements regarding the particle spectrum. However, by extrapolation and subtraction of the background and plotting the β -particle intensity logarithmically as before, an end-point at about 0.48 g./cm.^2 may be obtained. Thus an upper limit of energy at about $1.2 \times 10^6 \text{ eV}$ is indicated—though, until the precise form of the particle spectrum is more accurately known, the exact significance of this result is difficult to assess. At least it is not very different from $1.3 \times 10^6 \text{ eV}$, the value given by Pool (1938) from expansion chamber experiments. At a later stage we shall have to discuss whether the negative electrons constituting this energy spectrum represent a distinct mode of disintegration of the long-lived Ag^{106} ($\text{Ag}^{106} \rightarrow \text{Cd}^{106}$), as supposed by Pool, or whether they are secondary electrons emitted as a result of the process of transformation by K -electron capture which is supposed to occur. After a rough correction for absorption of the β -particles in the material of the source, a γ/β ratio of $1 : 1.05 \pm 0.1$ was deduced from fig. 6.

Attempts to detect the emission of palladium K -radiation from a thin piece of activated silver were made without success. Arranging for the effective absorption of the β -particles in 0.4 g./cm.^2 of graphite, no further absorption could with certainty be established—even with a xenon filled

counter—for suitably chosen thin foils either of molybdenum or of silver. The molybdenum foil ($\sim 13 \text{ mg./cm.}^2$)* should have reduced the intensity of any Pd K -radiation to about $1/3$, the silver foil (of about the same superficial density) only to $5/6$ of its original value. Failure to establish any difference in either case need not, however, be regarded seriously at this stage. The small intensity of the sources (total γ -ray effect roughly 30 counts per min.) and the low efficiency of the counters for the K -radiation is probably sufficient explanation. In spite of the negative result, then, we shall for the present accept the original suggestion without further experiment, assuming that the observed γ -radiation does in fact follow capture of a K -electron by the nucleus Ag^{106} . On this hypothesis, taking the results of our previous measurements of positron energy, the amount of energy

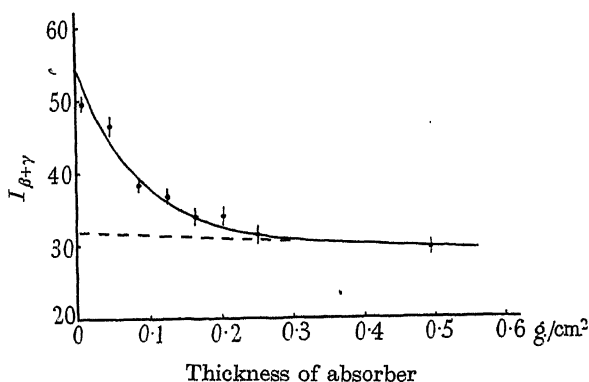


FIG. 6

available for radiation as quanta (or as kinetic energy of a neutrino) in this transformation is $(2.04 + 1.02 \pm \delta) \times 10^6 \text{ eV}$, $\delta \times 10^6 \text{ eV}$ being the assumed small difference in energy between the short- and long-lived isomeric forms of the nucleus Ag^{106} . This result, moreover, is quite independent of the mass of the neutrino, if this is regarded as finite.

Absorption measurements on the γ -radiation from the long-lived isomer were made after activation of much thicker pieces of silver (up to 1.1 g./cm.^2). Tungsten (to the extent of 13.9 g./cm.^2) and lead (up to 10 g./cm.^2) were used as absorbing materials. Over this range of thickness an effective quantum energy of $(0.60 \pm 0.03) \times 10^6 \text{ eV}$ was deduced from the measurements, in general agreement with the cloud chamber (recoil electron) results of Pool. On the basis of these results Pool (1938) has suggested

* We wish to thank Mr L. J. Davies, of British Thomson Houston Company, for sending us a small piece of this foil.

quanta of 0.3×10^6 to 1.0×10^6 eV energy. From the γ/β ratio obtained in our experimental arrangement we should estimate that roughly 100 quanta of effective quantum energy 0.6×10^6 eV are emitted for every negative electron of the supposed continuous spectrum.

When coincidence measurements on this γ -radiation were begun it was obvious at once that a very high relative coincidence rate was operative. As the experiment was continued, using different sources and different resolving times for the coincidence recording circuit, extreme values of $(1.80 \pm 0.09) \times 10^{-3}$ and $(2.23 \pm 0.09) \times 10^{-3}$ were obtained for the relative rate, R . A mean value of $(2.03 \pm 0.03) \times 10^{-3}$ depends upon a total of 982 min. counting with a source in position, in which time 1950 coincidences were recorded. On the basis of fig. 2, for a quantum of 0.6×10^6 eV energy, $(\alpha\epsilon) = 0.57 \times 10^{-3}$ —then equation (5) suggests the emission of 4 or 5 such quanta in cascade in order to explain the experimentally determined value of R . Even if, in actual fact, quanta of energies between 0.3×10^6 and 1.0×10^6 eV are involved, the conclusion cannot be very different: 4 or 5 such quanta must be emitted in succession from each nucleus which is transformed by K -electron capture.* If this is so, by far the greater part of the available energy, $(3.06 \pm \delta) \times 10^6$ eV, must be emitted as radiation; the most probable mode of electron capture in this case resulting in the production of the residual nucleus, Pd^{106} , in a high state of excitation (the assumed neutrino having quite small energy). Fig. 7 has been drawn conservatively, on the assumption of four equal quanta, with the entirely arbitrary choice† of the long-lived isomer as representing the ground state of the initial nucleus. For each isomer the energy state effective for K -electron capture is shown, in addition to the state involved in positron emission. It is very clearly consistent with the general picture provided by fig. 7 to suppose that a large difference in angular momentum characterizes the isomeric forms of the radioactive nucleus Ag^{106} , just as we have already concluded in the case UX_2 - UZ . In the present case, however, further knowledge concerning the energies and intensities of the γ -radiations will certainly be necessary before quantum numbers can, even provisionally, be assigned. Apart from the main fact of metastability or isomerism, the feature chiefly in need of explanation here is the non-occurrence—at least

* If only two quanta were emitted in succession in every transformation we should also expect $R \sim 2 \times 10^{-3}$ if each had 1.5×10^5 eV energy—but all the evidence is opposed to so high a value for the mean quantum energy of the radiation.

† A decision here may eventually be reached by studying the threshold energies for the reactions $\text{Pd}^{106} (\text{H}^1, n^1) \text{Ag}^{106}$ leading to the two isomeric forms of Ag^{106} , cf. Du Bridge, Barnes, Buck and Strain (1938).

to a first approximation*—of the sixteen energetically possible disintegration modes which are not shown as taking place on fig. 7. For the sake of comparison, the level scheme suggested by Pool and Campbell (1938) is given in fig. 8.

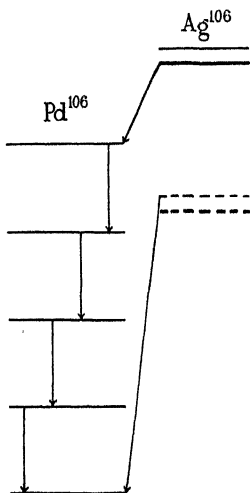


FIG. 7

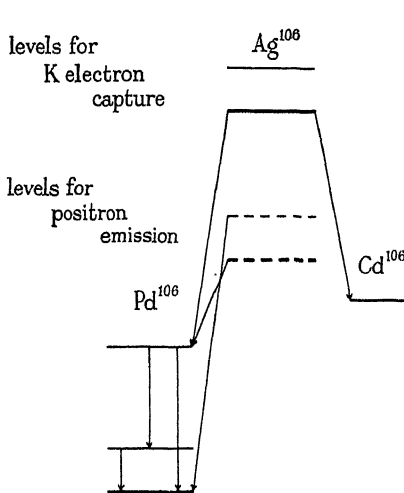


FIG. 8

Using a thin piece of activated silver, as in the β -particle absorption experiments, further coincidence studies were made with two tube counters carrying mica windows, arranged on the same axis, with the windows facing one another. The source was placed between the counters. Retaining this arrangement, the coincidence counting rate was determined with no additional absorbing material, with an absorber thick enough to cut out the effect of the β -particles on one side of the source only, and, finally, with such an absorber on each side of the source. A total of 1565 min. counting was carried out in this way and 565 coincidences were recorded. As the result of this work, time-correlation between β and γ counts, and a non-spurious effect amongst the β counts themselves, were clearly established. Evidently most of the negative electrons of the 8-day period are secondary to the K -electron capture process and do not represent an independent mode of disintegration of Ag^{106} . Regarding them as “internal conversion” electrons associated with the transitions schematized in fig. 7, our previous calcula-

* Pool and Campbell (1938) have reported the infrequent emission of positrons of small energy from the 8-day isomer. We have at present no information on this point—moreover our main conclusions are quite independent of it.

tion (p. 582) suggests a mean internal conversion coefficient of about 10^{-2} for the transitions in question. It may be remarked that Alvarez (1938) has already obtained evidence for internal conversion electrons from certain artificially produced radioelements of medium atomic number. In our case the energy spectrum of the electrons, whilst it does not appear clearly as a line spectrum from the absorption measurements, certainly covers the energy range to be expected.

Grants to one of us (J. V. D.) from the Board of Education and the Manchester Education Committee are gratefully acknowledged. As already indicated, the artificially radioactive sources which we used were prepared for us in the Cavendish High Voltage Laboratory. We wish to thank all whose work in development during the past year has brought this Laboratory into operation.

5. SUMMARY

The application of the method of coincidence counting by electrical means to the study of the time-correlation of nuclear processes—particularly that of successive quantum emission—is discussed in detail. It is shown that a knowledge of the energy dependence of the sensitivity of the counters is essential for any interpretation of the results of such experiments. For the γ -ray counters employed in the present investigation some information on this point has been obtained by the use of sources of thorium active deposit of short and long exposure. In the course of this work the general validity of the ideas of Ellis and Mott (1933) and of the nuclear level schemes of Oppenheimer (1936) has been established. Also the paired quanta of the annihilation radiation of positrons have been utilized in the calibration.

A study of the γ -radiations from UZ and the long-lived modification of Ag^{106} by the coincidence method has shown, in the former case, the emission of two quanta in succession in a very large fraction of the disintegrations, and, in the latter, the emission of 4 or 5 quanta in succession in the most probable mode. This result is in general accord with the view that nuclear isomers are to be distinguished one from the other by a small difference in energy content and a large difference in angular momentum. The particle emission from the short- and long-lived modifications of Ag^{106} has also been studied. It appears probable that the negative electrons of the long-period activity are internal conversion electrons associated with nuclear transitions following electron capture, rather than that they represent a distinct mode of disintegration of the long-lived radioelement, Ag^{106} .

REFERENCES

- Alichanian, Alichanow and Arzimovitch 1936 *Nature, Lond.*, **137**, 703-4.
 Alvarez 1938 *Phys. Rev.* **53**, 606.
 Aoki 1937 *Proc. Phys. Math. Soc. Japan*, **19**, 799-805.
 Bothe and v. Baeyer 1935 *Nachr. Ges. Wiss. Göttingen*, **1**, 195-7.
 Bothe and Gentner 1937 *Naturwissenschaften*, **25**, 90, 126.
 Cockcroft and Walton 1932 *Proc. Roy. Soc. A*, **137**, 229-42.
 v. Droste 1936 *Z. Phys.* **100**, 529-33.
 — 1937 *Z. Phys.* **104**, 474.
 Du Bridge, Barnes, Buck and Strain 1938 *Phys. Rev.* **53**, 447-53.
 Ellis 1938 *Proc. Phys. Soc.* **50**, 213-16.
 Ellis and Mott 1933 *Proc. Roy. Soc. A*, **141**, 502-11.
 Feather 1930 *Phys. Rev.* **35**, 1559-67.
 Feather and Bretscher 1938 *Proc. Roy. Soc. A*, **165**, 530-51.
 Feather and Dunworth 1938a *Proc. Camb. Phil. Soc.* **34**, 435-41.
 — — 1938b *Proc. Camb. Phil. Soc.* **34**, 442-9.
 Fleischmann 1936 *Z. Phys.* **103**, 113-24.
 Geiger and Marsden 1910 *Phys. Z.* **11**, 7-11.
 Heyn 1937 *Nature, Lond.*, **139**, 842.
 Klemperer 1934 *Proc. Camb. Phil. Soc.* **30**, 347-54.
 Krauss and Cork 1937 *Phys. Rev.* **52**, 763-8.
 Maier-Leibnitz 1936 *Z. Phys.* **101**, 478-85.
 Neher and Harper 1936 *Phys. Rev.* **49**, 940-3.
 Oppenheimer 1936 *Proc. Camb. Phil. Soc.* **32**, 328-35.
 Pool 1938 *Phys. Rev.* **53**, 116-23.
 Pool and Campbell 1938 *Phys. Rev.* **53**, 272-4.
 Pool, Cork and Thornton 1937 *Phys. Rev.* **52**, 380.
 Rossi 1930 *Nature, Lond.*, **125**, 636.
 Sizoo and Willemsen 1938 *Physica*, **5**, 105-10.
 Werner 1934 *Z. Phys.* **90**, 384-402; **92**, 705-27.
 Widdowson and Champion 1938 *Proc. Phys. Soc.* **50**, 185-95.
 Yukawa and Sakata 1937 *Sci. Pap. Inst. Phys. Chem. Res., Tokyo*, **31**, 187-94.

INDEX TO VOLUME CLXVIII (A)

- Age-hardening alloys, specific heat-temperature curves of (Swindells and Sykes), 237.
- Ammonium and ethylammonium cyanates, rates of transformation (Miller and Nicholson), 206.
- Andrade (E. N. da C.) *See* Bragg and others.
- Arley (N.) On the theory of coincidence experiments on cosmic rays, 519.
- Arnot (F. L.) and Beckett (Clark) A new process of negative-ion formation. IV, 103.
- β -particles, the scattering of (Barber and Champion), 159.
- Barber (A.) and Champion (F. C.) The scattering of fast β -particles, 159.
- Beckett (Clark) *See* Arnot and Beckett.
- Berg (W. F.) and Mendelssohn (K.) Photographic sensitivity and the reciprocity law at low temperatures, 168.
- Binnie (A. M.) The use of a vertical pipe as an overflow for a large tank, 219.
- Bragg (W. L.) and others. A discussion on plastic flow in metals, 302.
- Buckingham (R. A.) The classical equation of state of gaseous helium, neon and argon, 264.
- Buckingham (R. A.) *See* Massey and Buckingham.
- Burcham (W. E.) and Smith (C. L.) Experiments on the transmutation of fluorine by protons and deuterons, 76.
- Champion (F. C.) *See* Barber and Champion.
- Christopherson (D. G.), Gemant (A.), Hogg (A. H. A.) and Southwell (R. V.) Oscillatory motion of a fluid along a circular tube, 351.
- Christopherson (D. G.) and Southwell (R. V.) Relaxation methods applied to engineering problems. III, 317.
- Collens (H.) *See* Schonland and others.
- Coriolis perturbation in methane spectrum (Jahn), 469, 495.
- Cosmic rays, theory of coincidence experiments (Arley), 519.
- Dosch (C. H.) *See* Bragg and others.
- Diffusion velocity, porous diaphragm method (Hartley and Runnicles), 401.
- Discussion on plastic flow in metals, 302.
- Dunworth (J. V.) *See* Feather and Dunworth.
- Egerton (A. C.) *See* Harris and Egerton.
- Eisenschitz (R.) The specific heat of β -brass, 546.
- Electromagnetic energy of a point charge (Pryce), 389.
- Feather (N.) and Dunworth (J. V.) A further study of the problem of nuclear isomerism: the application of the method of coincidence counting to the investigation of the γ -rays emitted by uranium Z and the radioactive silver Ag^{106} , 566.
- Films on vitreous silica, thermal properties of (Palmer), 190.
- Forester (G. O.) *See* Tolansky and Forester.

Gamma rays emitted by uranium and silver (Feather and Dunworth), 566.

Harris (E. J.) and Egerton (A. C.) The decomposition and ignition of peroxides. I. Diethylperoxide, 1.

Hartley (G. S.) and Runnicles (D. F.) The determination of the size of paraffin-chain salt micelles from diffusion measurements, 420.

Hartley (G. S.) and Runnicles (D. F.) The porous diaphragm method of measuring diffusion velocity, and the velocity of diffusion of potassium chloride in water, 401.

Hatfield (W. H.) *See* Bragg and others.

Helium, low temperature properties (Massey and Buckingham), 378.

Helium, neon and argon, the classical equation of state of (Buckingham), 264.

Hilton (W. F.) Thermal effects on bodies in an air stream, 43.

Hinshelwood (C. N.) *See* Lewis and Hinshelwood.

Isotopes, separation of (Yates), 148.

Jahn (H. A.) A new Coriolis perturbation in the methane spectrum, 469, 495.

Lewis (R. M.) and Hinshelwood (C. N.) The thermal decomposition of nitrous oxide, 441.

Lightning, progressive (Schonland and others), 455.

Malan (D. J.) *See* Schonland and others.

Massey (H. S. W.) and Buckingham (R. A.) The low-temperature properties of gaseous helium, 378.

Mendelssohn (K.) *See* Berg and Mendelssohn.

Miller (C. C.) and Nicholson (J. R.) The rates of transformation in ethyl alcohol of ammonium and ethylammonium cyanates to the corresponding areas, 206.

Mott (N. F.) *See* Bragg and others.

Negative ions, formation of (Smith), 19.

Negative-ion formation, new process (Arnot and Beckett), 103.

Negative ions, formation by positive-ion impact on surfaces (Sloane and Press), 284.

Nicholson (J. R.) *See* Miller and Nicholson.

Nitrous oxide decomposition (Lewis and Hinshelwood), 441.

Nuclear isomerism of uranium and silver (Feather and Dunworth), 566.

Nuclear spin of iodine (Tolansky and Forester), 78.

Nuclear transmutations, investigation of (Yates), 148.

Orowan (E.) *See* Bragg and others.

Oscillatory motion of a fluid along a circular tube (Christopherson and others), 351.

Palmer (W. G.) The thermal properties and heats of adsorption of films on vitreous silica, 190.

Paraffin-chain salt micelles (Hartley and Runnicles), 420.

Peroxides, decomposition and ignition of (Harris and Egerton), 1.

- Photographic sensitivity at low temperatures (Berg and Mendelssohn), 168.
Plastic flow in metals, discussion, 302.
Press (R.) *See* Sloane and Press.
Preston (G. D.) *See* Bragg and others.
Pryce (M. H. L.) The electromagnetic energy of a point charge, 389.
- Reciprocity law at low temperatures (Berg and Mendelssohn), 168.
Relativistic optics, new system of normal co-ordinates (Temple), 122.
Relaxation methods applied to engineering problems, III (Christopherson and Southwell), 317.
Runnicles (D. F.) *See* Hartley and Runnicles.
- Schonland (B. F. J.), Malan (D. J.) and Collens (H.) Progressive lightning, VI, 455.
Sloane (R. H.) and Press (R.) The formation of negative ions by positive-ion impact on surfaces, 284.
Smith (C. L.) *See* Burcham and Smith.
Smith (R. A.) Formation of negative ions at metal surfaces, 19.
Southwell (R. V.) *See* Christopherson and Southwell *and* Christopherson and others.
Specific heat of β -brass (Eisenschitz), 546.
Specific heat-temperature curves of alloys (Swindells and Sykes), 237.
Superlattices, statistical theory of (Wang), 56, 68.
Swindells (N.) and Sykes (C.) Specific heat-temperature curves of some age-hardening alloys, 237.
Sykes (C.) *See* Swindells and Sykes.
- Taylor (G. I.) *See* Bragg and others.
Temple (G.) New systems of normal co-ordinates for relativistic optics, 122.
Thermal effects on bodies (Hilton), 43.
Tolansky (S.) and Forester (G. O.) The nuclear spin of iodine. III. Further measurements upon the fine structures in the first spark spectrum, 78.
Transmutation of fluorine, experiments on (Burcham and Smith), 176.
- Vertical pipe as overflow (Binnie), 219.
- Wang (J. S.) Statistical theory of superlattices with long-range interaction. I. General theory, 56.
Wang (J. S.) Statistical theory of superlattices with long-range interaction. II. The simple cubic lattice and the body-centred cubic lattice, 68.
- Yates (E. L.) The separation of isotopes for the investigation of nuclear transmutations, 148.